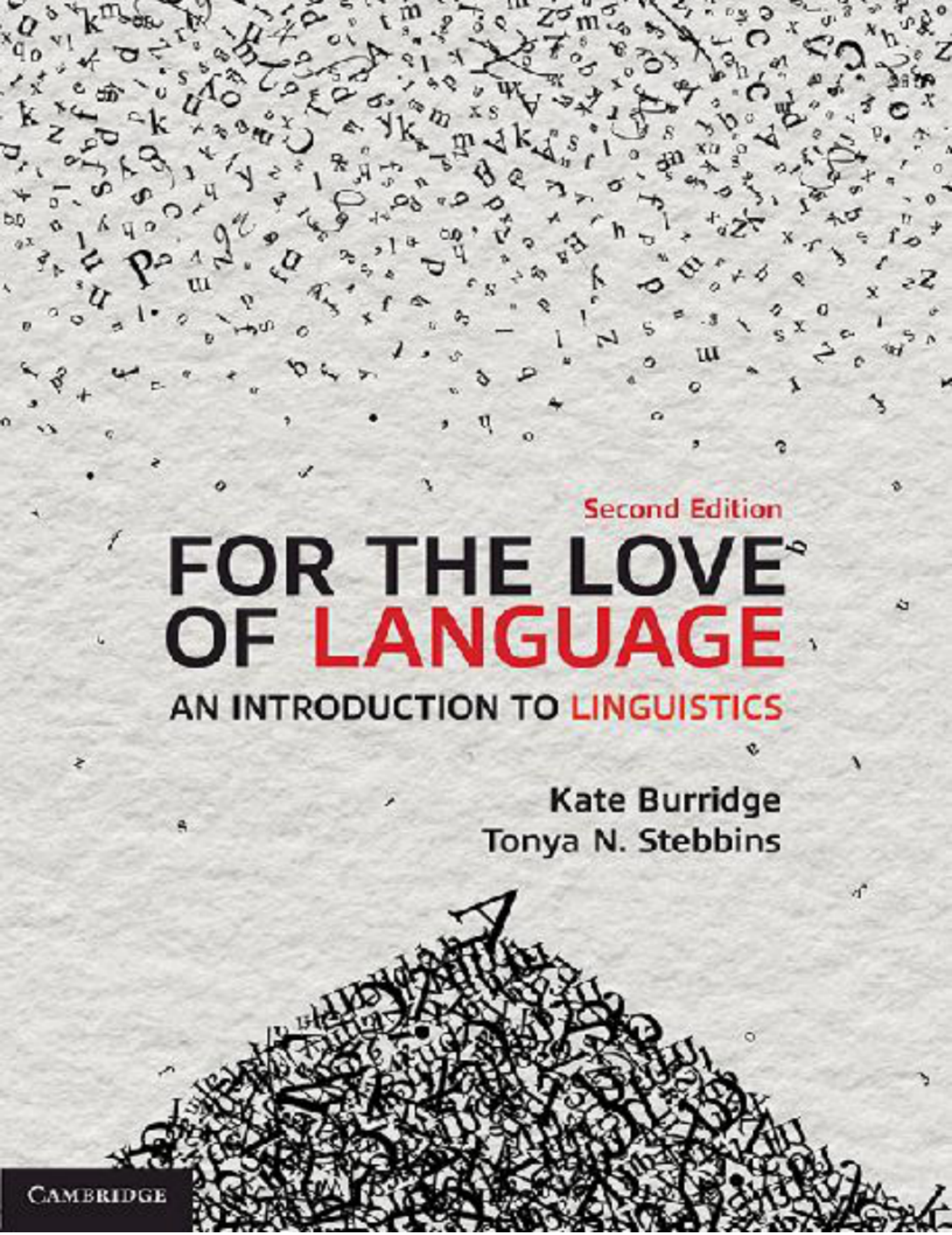


Second Edition

FOR THE LOVE OF LANGUAGE

AN INTRODUCTION TO LINGUISTICS

Kate Burridge
Tonya N. Stebbins

CAMBRIDGE

Second Edition

For The Love of Language
An Introduction to Linguistics

For the Love of Language: An Introduction to Linguistics is a comprehensive and engaging introduction to human language and the role of linguistics in understanding its fundamental design, acquisition and functions. Replete with case studies and examples from Australia, New Zealand and around the world, this text offers a thorough introduction to core topics, including the structure and meaning of words, the systems that organise language, strategies for learning about language, the evolution of language and the function of language as a complex social resource.

The second edition includes extensive new content across the entire text, including the areas of orthography, syntax, corpus linguistics, language acquisition and multilingualism. Each topic is accompanied by a wide array of pedagogical resources designed to consolidate student understanding, including examples, exercises and discussion points. Each chapter ends with a research project, providing readers with an opportunity to build on fundamental skills and engage more thoroughly with each topic through independent research.

Written by leading authors in the field of linguistics, *For the Love of Language* is an essential resource for undergraduate students in foundation linguistics.

Kate Burridge is Professor of Linguistics in the School of Languages, Cultures and Linguistics, Monash University, and a fellow of the Australian Academy of the Humanities.

Tonya N. Stebbins is Professor of Linguistics at La Trobe University.

Second Edition

For The Love of Language

An Introduction to Linguistics

Kate Burridge

Tonya N. Stebbins

CAMBRIDGE
UNIVERSITY PRESS

CAMBRIDGE
UNIVERSITY PRESS

University Printing House, Cambridge CB2 8BS, United Kingdom

One Liberty Plaza, 20th Floor, New York, NY 10006, USA

477 Williamstown Road, Port Melbourne, VIC 3207, Australia

314–321, 3rd Floor, Plot 3, Splendor Forum, Jasola District Centre,
New Delhi – 110025, India

79 Anson Road, #06–04/06, Singapore 079906

Cambridge University Press is part of the University of Cambridge.

It furthers the University's mission by disseminating knowledge in the pursuit of education, learning and research at the highest international levels of excellence.

www.cambridge.org

Information on this title: www.cambridge.org/9781108701013

© Cambridge University Press 2016, 2020

This publication is copyright. Subject to statutory exception and to the provisions of relevant collective licensing agreements, no reproduction of any part may take place without the written permission of Cambridge University Press.

First published 2016

Second edition 2020

Cover and text designed by Anne-Marie Reeves

Typeset by Integra Software Services Pvt. Ltd

Printed in Singapore by Markono Print Media Pte Ltd, October 2019

A catalogue record for this publication is available from the British Library

A catalogue record for this book is available from the National Library of Australia

ISBN 978-1-108-70101-3 Paperback

Additional resources for this publication at www.cambridge.edu.au/academic/linguistics

Reproduction and communication for educational purposes

The Australian *Copyright Act 1968* (the Act) allows a maximum of one chapter or 10% of the pages of this work, whichever is the greater, to be reproduced and/or communicated by any educational institution for its educational purposes provided that the educational institution (or the body that administers it) has given a remuneration notice to Copyright Agency Limited (CAL) under the Act.

For details of the CAL licence for educational institutions contact:

Copyright Agency Limited
Level 12, 66 Goulburn Street
Sydney NSW 2000
Telephone: (02) 9394 7600
Facsimile: (02) 9394 7601
E-mail: memberservices@copyright.com.au

Cambridge University Press has no responsibility for the persistence or accuracy of URLs for external or third-party internet websites referred to in this publication and does not guarantee that any content on such websites is, or will remain, accurate or appropriate.

Please be aware that this publication may contain several variations of Aboriginal and Torres Strait Islander terms and spellings; no disrespect is intended. Please note that the terms 'Indigenous Australians' and 'Aboriginal and Torres Strait Islander peoples' may be used interchangeably in this publication.

Contents

Preface

About the authors

Symbols used in this book

Acknowledgements

Using the online student resources

Using the online instructor resources

Part 1 ‘Using language to expose language’: Setting the scene

1 What is language?

- **1.0** Introduction
- **1.1** Hockett’s design features of language
- **1.2** Signs
- **1.3** Arbitrariness of signs
- **1.4** Language modalities
- **1.5** ‘Infinite use of finite means’: Rules and paradigms
- **1.6** Prescriptive and descriptive approaches
- **1.7** The ideal and the real
- **1.8** Language in its social context
- **1.9** Language across time
- **1.10** The case-study languages
    - Auslan
    - English
    - German
    - Gurindji
    - Japanese
    - Mali (Baining)
    - Māori
    - Russian

    - Turung
  - Further reading
  - Exercises and discussion points
  - Research project

- **2** What linguists do
  - **2.0** Introduction
  - **2.1** Everyone should know about language
  - **2.2** Linguistics at work
    - Linguists in the legal system
    - Linguists in education
    - Linguists and communication
    - Linguists in the field
    - Linguists in advertising and marketing
    - Linguists in health sciences
    - Linguists in business and professional services
    - Linguists in the creative arts
    - Linguists in computing and technology
  - **2.3** Some final remarks
  - Further reading
  - Exercises and discussion points
  - Research project

- **3** Ways to study language
  - **3.0** Introduction
  - **3.1** Sources of linguistic data
    - Corpora
    - Observational data
    - Interview data, questionnaires and focus groups
    - Experimental data
  - **3.2** Uses of linguistic data
    - Quantitative analysis
    - Qualitative analysis
  - **3.3** Ethics in linguistics
    - Human research ethics
    - Power, representation and authorship

    - Relationships and research partnerships
  - Further reading
  - Exercises and discussion points
  - Research project

Part 2 ‘In the beginning was the word’: Words, word structure and meaning

- **4** What’s in a word?
  - **4.0** Introduction
  - **4.1** Defining a word
    - Three facets of wordhood
  - **4.2** Lexical versus grammatical words
  - **4.3** Formulaic expressions (or lexical chunks)
  - **4.4** Magical words: Names
  - **4.5** Lexical addition
    - Acronyms
    - Shortenings
    - Compounding
    - Affixation
    - Backformation
    - Blends
    - Conversion
    - Commonisation
    - Reduplication
    - Borrowing
  - **4.6** Word mortality
  - Further reading
  - Exercises and discussion points
  - Research project

- **5** Morphology: The structure of words
  - **5.0** Introduction
  - **5.1** Words, lexemes and morphemes
  - **5.2** Roots, stems and affixes

- 5.3 Inflection and derivation
  - Paradigms
- 5.4 Morphs and allomorphs
- 5.5 Lexical and grammatical morphemes
- 5.6 Word classes and parts of speech
  - Nouns
  - Verbs
  - Adjectives and adverbs
- 5.7 More morphological processes
  - Modification
  - Suppletion
- 5.8 Productive morphemes
- Further reading
- Exercises and discussion points
- Research project

6 Semantics: The meaning of words
- 6.0 Introduction
- 6.1 The meaning of meaning
- 6.2 Meaning properties: Ambiguity and anomaly
- 6.3 Lexical relations – all those ‘nyms’
  - Synonymy
  - Antonomy
  - Homonymy
  - Polysemy
  - Hyponymy
- 6.4 Meaning components
- 6.5 Meaning and context
  - ‘You shall know a word by the company it keeps’: Collocation
  - Idioms: Special examples of collocation
- 6.6 Metonymy and metaphor
- 6.7 Semantic change: Types of meaning shifts
  - Broadening
  - Narrowing
  - Shifting

- Changing values: Elevation and deterioration
- Why words shift their meanings

6.8 Slips of the ear and brain

6.9 Phonesthemes

Further reading

Exercises and discussion points

Research project

Part 3 ‘The deep grooves of language’: Sounds and grammar

7 Phonetics

7.0 Introduction

7.1 Articulatory phonetics and the vocal tract

7.2 Consonants

- Voicing
- Manner of articulation
  - Stops
- Fricatives
- Affricates
- Approximants
- The International Phonetic Alphabet

7.3 Vowels

- The basics
- Australian and New Zealand English
- Other languages
- Diphthongs

7.4 Uncomfortable relations: The written and spoken word

- Sounds to symbols

7.5 Prosody

- Syllables
- Pitch and intonation
- Tone
- Paralinguistic features

7.6 Unpacking the parts
    More mis-analysis: The mondegreen
    Phonetics and sound symbolism
7.7 Acoustic phonetics
    Visualising sound
    Acoustic analysis: How it is used
    Forensic phonetics
    Sociophonetics
    Sound change
Further reading
Exercises and discussion points
Research project

8 Phonology: The sound system
8.0 Introduction
8.1 Phonemes, phones – and allophones
    Why bother with allophones and phonemes?
8.2 Sounds in connected speech: Phonological processes
    Changes that alter syllable structure: deletion, addition and reordering of sounds
    Processes that involve the modification of sounds
8.3 Determining phonemes
    Minimal pairs
    Complementary distribution
    So-called ‘free variation’
    Phonetic similarity
    Metaphors to help illuminate the phoneme
8.4 Distinctive features and natural classes
8.5 The pronunciation of morphemes: Morphs and allomorphs
8.6 Prosodic phonology
    Syllable structure
    Phonotactic constraints
    Prosodic features
8.7 Representing spoken language
    The development of writing

- Further reading
- Exercises and discussion points
- Research project

9Syntax: The structure of sentences
- **9.0**Introduction
- **9.1**Grammaticality
- **9.2**The basic clause
- **9.3**Sentence types (simple, coordinate, complex)
- **9.4**Constituency and phrases
- **9.5**Categories and functions
- **9.6**Grammatical relations and semantic roles
  - Syntactic ambiguity
- **9.7**Phrase structure rules
  - Recursion
- **9.8**Word order types
- **9.9**Transitivity
  - The subject
  - The object
  - Predicative complements
  - Adverbials
- **9.10**Sentence types
- **9.11**Some final thoughts
- Further reading
- Exercises and discussion points
- Research project

Part 4‘Language is a social fact’: Variation and change

10Variation and identity
- **10.0**Introduction
- **10.1**‘The chaos of a living speech’
- **10.2**Languages, dialects, accents – and standards
  - Standardised language

- When do dialects become languages?
- **10.3** User-related differences: Regional variation
  - Case study: Variation in Australian and New Zealand English
  - Case study: Regional variation in Auslan
- **10.4** User-related differences: Social variation
  - Case study: Age-related variation
  - Case study: Ethnicity-related variation
- **10.5** Use-related differences: Situational variation
- **10.6** Style: What makes individuals and groups distinctive?
  - Slanguage and swearing
- **10.7** Online identities
- **10.8** The ‘chaos of a living speech’ is not so chaotic
  - Researching language variation
  - Developing a project
- Further reading
- Exercises and discussion points
- Research project

11 Structure across time
- **11.0** Introduction
- **11.1** The ‘what’ of change: Phonology
  - Old sounds drop out: Deletion
  - New sounds appear: Epenthesis
  - Existing sounds are modified: Assimilation, dissimilation and metathesis
  - Phonetic versus phonological change
- **11.2** The ‘what’ of change: Grammar
  - Change in morphology: Morphological types
  - Change in morphology: Analogy
  - Change in syntax: Word order types
  - Change in syntax: Creating grammar
- **11.3** The ‘how’ of change
  - All changes occur in a social context
  - Changes proceed gradually through the system
- **11.4** The ‘why’ of change

- Physiological factors: Simplicity
- Social factors: The speech community
- Frequency factors: The effects of repetition

11.5 Language families and establishing a genetic relationship
- How to confirm genetic relationships
- How do language families arise?
- Mapping change as trees and waves

11.6 Attitudes to change

Further reading

Exercises and discussion points

Research project

12 Languages and cultures in contact

12.0 Introduction

12.1 Types of contact between languages
- The ‘monolingual mindset’
- The outcomes of language contact

12.2 Language maintenance
- Diglossia
- Translanguaging
- Code-switching
- Structural borrowing
- Areal linguistics

12.3 Language shift
- Contact-induced language change: Long-term patterns of shift
- Language maintenance and reclamation
- Factors contributing to language shift

12.4 The creation of new languages
- Pidgin languages
- Creoles
- New mixed languages

12.5 Language choice and linguistic accommodation

12.6 Translation and interpreting

Further reading

- Exercises and discussion points
- Research project

Part 5‘Language that rolls up its sleeves’: Language at work

13Language, context and interaction
  - **13.0**Introduction
  - **13.1**Information packaging
    - Given information comes before new
    - Topic comes before comment
    - Principle of front focus
  - **13.2**Special discourse strategies
    - Passives: Creating new subjects
    - Existentials (or there-constructions)
    - Extraposition
    - Cleft constructions
    - Fronting
    - Dislocation
  - **13.3**Cohesion
    - Reflexive pronouns
    - Deixis: ‘Pointing words’
  - **13.4**Discourse analysis
    - Big D and little d discourse analysis
    - Discourse markers
  - **13.5**Different types of communication
  - **13.6**Conversations
    - Expressing support
    - Turn-taking
  - **13.7**Narrative structures
    - Cultural aspects of narratives
  - **13.8**Ethnography of communication
  - Further reading
  - Exercises and discussion points

Research project

14 Language, context and understanding
  14.0 Introduction
  14.1 Some preliminary observations
    The conduit metaphor
  14.2 Context: Linguistic and situational
  14.3 The force of words
  14.4 The cooperative principle
    Maxim of quantity
    Maxim of quality
    Maxim of relation
    Maxim of manner
  14.5 Extracting meaning from talk
    Contravening maxims
    Presuppositions
    Inferencing
    Implicature
  14.6 (Im)politeness
    Face
  14.7 Cross-cultural pragmatics
    Silence across cultures
    When insults aren’t insults: The ‘banter principle’
  14.8 Gesture in communication
    Teasing apart sign language and gesture
  Further reading
  Exercises and discussion points
  Research project

15 Language and social values
  15.0 Introduction
  15.1 Linguistic prejudice and stereotyping
  15.2 The identity functions of language
  15.3 Verbal hygiene: Doctrines of correctness and linguistic purism
    Style guides

    - Good grammar
    - Political correctness
    - Language and gender
  - **15.4** Prescription, competence and social inclusion
    - Educational implications of non-standard variation
  - **15.5** Language planning
    - Language planning: Who decides?
    - Graphisation
    - Modernisation
    - Standardisation
  - **15.6** Language and nation
    - Ideologies of national language planning
  - **15.7** Research on language ideologies in practice
  - **15.8** Final thoughts
  - Further reading
  - Exercises and discussion points
  - Research project

Part 6 ‘Language is the dress of thought’: Language, mind and world

- **16** Language, the mind and the brain
  - **16.0** Introduction
  - **16.1** Language, culture and the mind
  - **16.2** How language maps onto the brain
  - **16.3** Ways of learning about language and the brain
    - Behavioural methods of psycholinguistic research
    - Neuroimaging techniques
  - **16.4** Models of language processing
  - **16.5** Language production
    - Error analysis
    - Lexical retrieval
  - **16.6** Comprehending words
  - **16.7** Comprehending sentences

- 16.8 The bilingual brain
- 16.9 Language disability
  - Developmental disability
  - Acquired disability
- Further reading
- Exercises and discussion points
- Research project

17 Language acquisition
- 17.0 Introduction
- 17.1 Approaches to language acquisition research
- 17.2 Language acquisition pathways
  - A note about ‘input’
  - The critical period hypothesis
- 17.3 Acquisition of phonology
  - The fis phenomenon
- 17.4 Acquisition of vocabulary
- 17.5 Acquisition of grammar
  - Analogy at work: Wugs that bing
  - Putting words together
- 17.6 Other things you need to know to be a user of language
  - Acquiring honorifics in Japanese
- 17.7 Development of literacy
  - Exploration of a literate identity
  - Exploration of text features
  - Early reading
  - Reading for meaning
  - Independent reading
  - Developing writing
- 17.8 Second-language acquisition
  - Developmental patterns in second-language acquisition
  - The impact of the first language in second-language acquisition
- Success factors in second-language acquisition
- 17.9 Growing up bilingual

- Acquisition of phonology
- Acquisition of vocabulary
- Acquisition of grammar
- Language awareness and differentiation
- The benefits of growing up bilingual

Further reading
Exercises and discussion points
Research project

18 Computational linguistics
Simon Musgrave

18.0 Introduction
18.1 How does HAL understand what is said?
- Speech recognition
- Syntactic and semantic analysis

18.2 How does HAL say anything?
18.3 Machine translation
18.4 Corpus linguistics
18.5 Techniques from genetics
18.6 Conclusion
Further reading
Exercises and discussion points
Research project

References
Index

Preface

After many years of teaching first-year linguistics from a variety of textbooks, piles of notes, selected readings and so on, we decided to write the textbook we had always wanted – a textbook that:

- covered a full year of content to save on textbook costs for students
- included comprehensive exercises, discussion points, ideas for research projects to facilitate team teaching and course planning
- showcased the applications of linguistics to real life
- focused on a small number of 'case-study languages' so students could get a deeper sense of language diversity
- was accessible but provided strong foundations for later years. Each of these choices has had an influence on the book you are now reading – and challenged us to think about new ways to juggle the competing priorities we set ourselves. Following is a brief summary of how the book is set up as a result.

Content

The text is arranged into six parts:

- Part I – 'Using language to expose language': Setting the scene
- Part II – 'In the beginning was the word': Words, their structure and meaning
- Part III – 'The deep grooves of language': Sounds and grammar
- Part IV – 'Language is a social fact': Variation and change
- Part V – 'Language that rolls up its sleeves': Language at work
- Part VI – 'Language is the dress of thought': Language, mind and world.

We imagine that two or three parts of the text would be enough to fill a twelve- to thirteen-week semester. As a result, there may be sections of the text that students do not cover in class. These could be set as pre-reading for later courses.

We chose this breadth of content to provide lecturers with flexibility regarding what they covered and also to demonstrate to students the relevance that linguistics has to so many areas of life. Naturally, if a unit focuses quite deeply on a specific topic, an additional reading or two may be needed.

Supporting materials

The book is best read and used in conjunction with the supporting materials available on the website. These include additional exercises and research projects as well as more information on the case-study languages and referencing conventions.

Case-study languages

Why case-study languages? And why these particular choices? Our case-study languages are Auslan, English, German, Gurindji (and Gurindji Kriol), Japanese, Mali, Māori, Russian and Turung. Each of these languages is introduced to readers on the website. We wanted to focus on a smaller number of languages than included in the typical introductory text so that students could gain familiarity with new languages and see how methods and concepts from linguistics aid this type of learning. We included English as a starting point in many chapters since, in our experience, it can be helpful to start students off with what they already know. It was also important to us to include Auslan, Māori and an Australian Aboriginal language because some awareness of these languages is important in many of the professions in which linguistics students find themselves working later in life. Other people may have made different choices – with around 6000 languages to choose from, there are lots of interesting options!

Style and referencing

We have used the friendly writing style we associate with popular science writing in this text and tried to write in the way we sound when we give lectures. This approach increases accessibility for readers new to linguistics. Consistent with this approach, we have avoided in-text referencing. Obviously this is not the model that students should follow in writing essays. In order to demonstrate the difference, we provide guidelines for presenting work, as well as an extract from our book with appropriate academic referencing and supporting discussion, in the online materials.

Linguistics matters

In this book, we seek to expose the wondrous workings of language – as well as we can in eighteen short chapters. At the same time, we want to highlight many of the ways in which people's lives are affected by language. We trust it will be a comfort to students to know that the things they learn in this book will prove useful. In particular, we want to show what can go wrong when people don't know enough about language and how it actually works.

Throughout the book, we address questions to do with: beliefs about class and dialect; bilingualism and language proficiency; correctness and incorrectness; manipulation through advertising and propaganda; the influences of language on thinking and behaviour; linguistic discrimination; and the maintenance of power. When a radio talkback caller condemns another for a dropped consonant or an aberrant apostrophe, the comment appears harmless enough. But the 'commonsense' beliefs people hold about their language often inform decisions that affect the life chances of others, and the consequences can be catastrophic.

Sociolinguist Diana Eades describes a significant legal case, a murder trial, where the oral evidence and reports of four linguists were central to shaping the judge's final decision. Students can read the judgment (*Western Australia v Gibson 2014)* at http://www.austlii.edu.au/au/cases/wa/WASC/2014/240.html. Because it is a long and difficult read (Diana Eade's report alone was 10 000 words!), Diana has kindly summarised the case for us in a piece that appears on the website: 'The Relevance of Linguistics to Human Rights in Police Interviews'.

Many of our colleagues, including Diana, have made it their life work to ensure that well-researched principles of linguistics are put above entrenched but inaccurate notions about how people speak (or should speak). Now, due to their hard work, linguistic evidence is playing a crucial role in dispensing justice in many important areas, such as the law, education and immigration. We wish we could do more in this book to flag this important work.

About the authors

Kate Burridge is Professor of Linguistics in the School of Languages, Cultures and Linguistics, Monash University, and a fellow of the Australian Academy of the Humanities. Her main areas of research are grammatical change in Germanic languages, the Pennsylvania German spoken by Amish/Mennonite communities in North America, the notion of linguistic taboo, popular perceptions about language, and the structure and history of English. She is a regular presenter of language segments on radio and TV.

Tonya N. Stebbins is Professor of Linguistics at La Trobe University. Her contributions to the field include innovative partnerships with communities who speak minority and endangered languages, and ensuring that communities have access to resources that support language revitalisation. Stebbins is an experienced consultant who has completed projects across the health, human services and education sectors. Projects include program evaluation, organisational development, program review, and qualitative research on topics such as family violence, child protection, alcohol and other drugs (AOD) interventions, homelessness and financial capability.

Symbols used in this book

Although linguistics relies heavily on abbreviations in a number of areas, we have tried to spell out most terms in this book. Where abbreviations are used, they are always introduced first in a spelt-out form. If the reader does come across an abbreviation that is not clear, the Leipzig Glossing Rules includes a list of standard abbreviations that provides a useful reference point (see https://www.eva.mpg.de/lingua/resources/glossing-rules.php).

The reader may also find it useful to refer to the materials on the inside cover of the book that set out the International Phonetic Alphabet (IPA), and provide references for phonemes and parts of speech in English.Other symbols that we use in the book but do not always define are listed below.

[ ] = phonetic representation (in IPA – see Chapters 7 and 8)

/ / = phonemic representation (in IPA – see Chapters 7 and 8)

< > = orthographic representation (see Chapters 7 and 8)

> = historical changes (see Chapter 11)

→ = processes (see Chapter 8)

$ = syllable boundary (see Chapters 7 and 8)

. = phonological word break within grammatical word

- = affix

= = clitic

* = ungrammatical sentence (see Chapter 9); and reconstructed form in a proto-language (see Chapter 11)

Acknowledgements

Needless to say, a book like this has depended upon the generous support of many people. First, we owe a huge debt of gratitude to those colleagues who generously supplied us with material and advice on various case-study languages (Felicity Meakins, Stephen Morey, Adam Schembri, Liz Pearce, Wes Robertson and Ian Cummins). We are grateful to Sarah Fishlock, Emma Murphy and Christina Eira, who bravely read draft chapters of the book – their advice and suggestions were invaluable. Special thanks must go to Simon Musgrave for his chapter on computational linguistics and to Jenny Price for her chapter on phonetics – their contributions have been spectacular. There are others we need to thank too. We thank our dear and tolerant friends, colleagues and students who have been so supportive during the period of writing this book, and, of course, we give our love and thanks to our wonderful families for their endless encouragement and generous patience.

For this second edition, we have additional thanks. A number of people have very generously allowed us to use their images. We are indebted to: Wes Robertson (for the railway signage), Headline Shirts (for the ümlaut T-shirt), Bob Dixon (for the 1910s Aussie slang clipping), Carolin Biewer (for the Hong Kong street scene), and Louisa Willoughby and Ramas McRae (for the Auslan images). We would also like to acknowledge the original contribution that Jenny Price made to Chapter 7 (on phonetics); this chapter could not have happened without her expertise. We are also grateful to Marko Pavlyshyn for his input to the Russian examples and to Liz Pearce and Miriam Meyerhoff for their careful reading of our Māori examples. We owe a huge debt of thanks to the students at La Trobe University, Monash University and Victoria University of Wellington – their eagle eyes spotted a number of errors in the first edition. Another huge thanks to Allie Severin for her insightful and meticulous reading of the manuscript. Finally, both of us are extremely grateful to our editors at Cambridge University Press over the years; we give our thanks in particular to Olivia Tolich and more recently Tanya Bastrakova.

Artwork

We are grateful to the following individuals and organisations for permission to use their material in *For the Love of Language*.

Figure 1.1: © Getty Images/Vectorios2016; **1.2**, **10.1**: Adapted from T. Johnston & A. Schembril (2007) *Australian Sign Language (Auslan): An Introduction to Sign Language Linguistics*. Cambridge: Cambridge University Press. Reproduced with permission; **1.3**: © Getty Images/calvindexter; **1.4**: Reproduced with permission from Louisa Willoughby, Monash University. Photos of Ramas McRea used with permission; **1.5**: © Getty Images/Juanmonino; **1.6**: Courtesy of Brenda L. Croft, 2014; **1.7**, **8.7**: Courtesy of Wes Robertson; **1.8**: © Getty Images/Mark Meredith; **1.9**: mytourapp.com/Wikimedia Commons. Reproduced under Creative Commons Attribution 3.0 Unported license; **1.10**: Courtesy of Stephen Morey; **3.1**, **3.2**: Adapted from R. Sigley & J. Holmes (2002) 'Looking at girls in corpora of English', *Journal of English Linguistics*, 30(2): 138–57; **6.1**: © Getty Images/stockakia; **6.2**: © Getty Images/erhui1979; **6.3**: (A) © Getty Images/isaxar, (B) © Getty Images/musmellow, (C) © Getty Images/Johner Images; **6.4**: (A) © Getty Images/arthobbit, (B) © Getty Images/arcady_31, (C) © Getty Images/Devilkae, (D) © Getty Images/KOHb, (E) © Getty Images/Tetra Images, (F) © Getty Images/ImageMediaGroup; **6.5**: © Getty Images/AZaytsev; **7.1**–**7.7**: R-A. Knight (2012) *Phonetics: A Coursebook*, Cambridge: Cambridge University Press. © Rachael-Anne Knight 2012. Reproduced with permission of Camridge University Press; **7.8**: IPA Chart, http://www.internationalphoneticassociation.org/content/ipa-chart, available under a Creative Commons Attribution-Sharealike 3.0 Unported License. Copyright © 2015 International Phonetic Association; **7.10**: IPA Chart, http://www.internationalphoneticassociation.org/content/ipa-vowels, available under a Creative Commons Attribution-Sharealike 3.0 Unported License. Copyright © 2005 International Phonetic Association; **7.12a**, **7.14**: F. Cox & S. Palethorpe (2007), 'Illustrations of the IPA: Australian English', *Journal of the International Phonetic Association*, 37(3): 341–50. Cambridge University Press. https://doi.org/10.1017.S0025100307003192. Reproduced with permission; **7.12b**, **7.15**: Adapted from Figure 1, p. 98 in L. Bauer, P. Warren, D. Bardsley, M. Kennedy and G. Major (2007) 'Illustrations of the IPA: New Zealand English', *Journal of the International Phonetic Association*, 37(1): 97–102. Cambridge University Press. https://doi.org/10.1017/S0025100306002383. Reproduced with permission; **7.13**: © Judy Horacek; **7.17**–**7.20**: Screenshots of Praat software reproduced with permission of the authors from Paul Boersma & David Weenink (2019) Praat: doing phonetics by computer [Computer program]. Available, retrieved 26 March 2019 from http://www.praat.org; **7.22**: (A) ©

Getty Images/Westend61, (B) © Getty Images/John Carnemolla; **8.6**: (a, c) © Getty Images/RichVintage, (b) © Getty Images/albertogagna; **8.8**: Dbachmann/Wikimedia Commons. Licensed under Attribution-ShareAlike 3.0 Unported; **10.2**: © Getty Images/Diane Labombarbe; **11.1**: Adapted from the IPA Chart, http://www.internationalphoneticassociation.org/content/ipa-chart, available under a Creative Commons Attribituion-Sharealike 3.0 Unported License. Copyright © 2005 International Phonetic Association; **11.2**: © Headline Shirts. This image has been reproduced with permission. This shirt (design and likeness) is the property of Headline Shirts and should not be reproduced without permission; **11.5**: © Getty Images/Rob Evans; **11.7**: © Getty Images/mathisworks; **12.1**: Courtesy of Carolin Biewer; **14.1**: (A) © Getty Images/Tzogia Kappatou, (B) © Getty Images/Wichai Treethidtaphat/EyeEm, (C) © Getty Images/Chatmali; **14.2**: © Getty Images/kornnphoto; **14.3**: © Getty Images/pixelliebe; **16.1**: © Getty Images/CSA Images; **16.2**: Blausen.com staff (2014). ‘Medical gallery of Blausen Medical 2014’. *Wiki Journal of Medicine* 1(2). DOI: 10.15347/wmj/2014. 010. Used under Creative Commons Attribution 3.0 Unported licence. **16.3**: © Getty Images/mustafahacalaki; **16.4**: © Getty Images/wetcake; **16.5**: H.J. Bigelow (1850) ‘Dr Harlow’s case of Recovery from the passage of an Iron Bar through the Head’, *The American Journal of Medical Sciences*, 20: 13–22; **16.9**: © Getty Images/cako74; **16.10**: Reproduced from P. Warren (2012) *Introducing Psycholinguistics*. Cambridge University Press. © Paul Warren 2013. Reproduced with permission of Cambridge University Press; **16.11**: F. Meakins et al. (2014) *Gurindji to English Dictionary*. Batchelor: Batchelor Press. Reproduced with permission of the authors and Batchelor Institute Press; **18.6**, **18.7**: Adapted from A. Calude (2017) ‘The use of heaps as quantifier and intensifier in New Zealand English’, *English Language and Linguistics*, pp. 1–26. Cambridge University Press. https://doi.org/10.1017/S1360674317000521. © Cambridge University Press 2017. Reproduced with permission of Cambridge University Press.

Text

Extracts from the Superlingo blog (https://www.superlinguo.com) in chapter 2 all © Lauren Gawne. Reproduced with permission of Lauren Gawne and interviewees; Material reproduced and adapted from K. Börjars & K. Burridge (2019). *Introducing English Grammar*, 3rd Edition, Routledge © 2019 Kersti Börjars and Kate Burridge. Reproduced with permission of The Licensor, Informa UK Limited, through PLSclear; Extracts reproduced from Burridge, K. (2010).

‘Linguistic cleanliness is next to godliness: taboo and purism’ *English Today*, 26(2), pp. 3–13. Cambridge University Press. https://doi.org/10.1017/S0266078410000027, © Cambridge University Press 2010, reproduced with permission; Text extract in ‘Linguistics and service provision for the Deaf’ reproduced with permission of Louise Willoughby; Extract from ‘Identifying, Teaching and Assessing Key Skills in Linguistics’ workshop reproduced with permission of Richard Hudson; Extracts from ‘Experiences of an outback linguist’ reproduced with permission of Eugenie Collyer; Extracts from ‘“Zombie rule that I don’t live by”: measuring language attitudes with a changing yardstick’ reproduced with permission of Allie Severin; Extracts from K. Burridge (2012). ‘Euphemism and language change’, Lexis, *E-Journal in English Lexicology*, © Lexis 2012, reproduced with permission; Extracts from K. Allan & K. Burridge (2006). *Forbidden Words: Taboo and the Censoring of Language*. Cambridge: Cambridge University Press, © 2006 Keith Allan and Kate Burridge, reproduced with permission of the authors and Cambridge University Press; Material reproduced and adapted from K. Burridge (2006). *Weeds in the Garden of Words: Further Observations on the Tangled History of the English Language*, Cambridge University Press © Kate Burridge 2004, 2005, reproduced with permission of the author and Cambridge University Press; Material reproduced and adapted from K. Burridge & B. Kortmann (2008). Introduction: varieties of English in the Pacific and Australasia. *The Pacific and Australasia* (pp. 23–38). https://doi.org/10.1515/9783110208412.0.23 © 2008 DE GRUYTER MOUTON. Reproduced with permission; Extract from M. Zijlstra (Presenter). (2009, May 30). ‘What we lose when languages disappear’: Interview with Nicholas Evans [Radio broadcast transcript]. In Linguafranca. Melbourne, Australia: Radio National, ABC. https://www.abc.net.au/radionational/programs/archived/linguafranca/dying-words/3133192, with permission of Nick Evans; Extracts from Little Red Riding Hood reproduced with permission of James Finn Garner ©1994; Transcription of basketball player story reproduced with permission of Tania Strahan; Extract from New Zealand Parliament Hansard Question Time, 28 October 2014: Source: Office of the Clerk/Parliamentary Service. Licensed by the Clerk of the House of Representatives and/or the Parliamentary Corporation on behalf of Parliamentary Service for re-use under the Creative Commons Attribution 4.0 International licence. Full licence available at https://creativecommons.org/licenses/by/4.0/.

Every effort has been made to trace and acknowledge copyright. The publisher apologises for any accidental infringement and welcomes information that would redress the situation.

Using the online student resources

A variety of student resources to support learning are provided on the companion website to this text at https://www.cambridge.edu.au/academic/linguistics.

You will come across this icon in the margin throughout the book, indicating that a useful video, resource or weblink is available on the companion website. Some items are tied to specific exercises in the book, while others are purely reference material to expand upon the content covered.

Essential concepts

Here we provide concise definitions for each term. This section will be useful for revision and may also help you to clarify your understanding as you read.

For a more extensive glossary of linguistics terminology, we recommend students use a dictionary of linguistics or an online glossary, such as the Online Dictionary of Language Terminology (ODLT).

Responses to exercises and discussion points in the book

This section provides a range of additional prompts, and describes the responses to the exercises and discussion points in each chapter.

Additional exercises, discussion points and

research projects

The best way to learn the new analytical skills and concepts associated with linguistics is through practice and reflection. An additional five exercises or discussion points for each chapter have been provided online.

In addition to the research projects suggested in the book (one for each chapter), there are two additional research projects per chapter online.

Research project tips and resources

Research projects can be a fantastic opportunity to get your hands dirty with linguistics data and analysis, but sometimes it can be hard to know quite where to start. The online materials provide additional guidance for each research project, including links and references that will get you started.

Further Links

The further links section provides a list of activated links, including links to further resources of interest, and websites that will help you to complete the exercises or discussion points and research projects in the chapters.

Further resources

The further resources section provides materials to extend your thinking in fun and interesting ways. Each further resources section has numbered subheadings; these numbers correspond to the section of the book that relates to the suggested resources.

Podcasts

Each part commences with a short podcast in which the authors introduce and discuss relevant and interesting topics in the field of linguistics.

Language biographies

In Chapter 1, you are introduced to the nine case-study languages discussed throughout the book. The purpose of this online material is to provide an extended introduction to each language and to signal where each language is of particular interest in the book.

Guide to academic referencing

This text follows the conventions of popular science writing and avoids detailed referencing. This is done in order to make the text as readable as possible. The online guide presents additional material to help you see how to use references, quotes and examples from other sources in your assessments.

IPA Flashcards

Phonetics are discussed in Chapter 7, but the IPA is discussed and used throughout the text. You will find the IPA on the inside back cover of the text. On the website, you will find printable flashcards that will enable you to revise and check your understanding of the IPA symbols.

Using the online instructor resources

A variety of resources for instructors are provided on the companion website to this text at https://www.cambridge.edu.au/academic/linguistics. These materials are designed to help instructors prepare lectures, tutorials and interactive class activities or assessments.

Quizzes

A set of ten extra questions is available to instructors. These can be set as assessment tasks or used in other class activities to check student's learning and progress.

PowerPoint Presentations

A set of PowerPoint slides for each chapter provides an outline for weekly lecture presentations and covers key concepts.

Part 1

◆

‘Using language to expose language’

Setting the scene

1 What is language?

Language is the most massive and inclusive art we know, a mountainous and anonymous work of unconscious generations.

Edward Sapir (1921), Language: An Introduction to the Study of Speech

1.0 Introduction

Are you a lover of language? Do you revel in solving crosswords? Are you picky about punctuation? Does the history of a word resonate for you? Do you roll sounds around in your mouth like mints? Are you the Scrabble champion of your social network? Do you pun your friends and colleagues under the table? If these descriptions fit you in any way, then you are probably going to find this book revelatory and inspiring because in these pages we map out the means for you to dive even deeper into the experience of being a language user and give you the tools to explore language in ever finer detail.

If you're not obsessively enamoured with language (and one of the authors isn't, going by the list of likes and aptitudes listed above), but you are interested in how it works and how it can work for you, then this book is for you too, since it gives you a platform from which you can view language in all its variety, providing a series of avenues for the exploration of its inner workings.

Language is a very significant part of what makes us human and, more particularly, it is what allows us to communicate with others as social beings. Language makes our social world go round and, for this reason, it is worth knowing a few things about how it works. There are also some really practical advantages to knowing about language; we outline some of these in Chapter 2.

In this chapter, we are concerned mainly with sorting out what language is. In order to map out this space, we ask the following questions.

- How do we 'do' language?
- What makes human language distinct from forms of animal communication?

- What are the special features of language?
- Where did language come from?
- How is it best to view language?
- What can we know about language?
- Where do we start?

The following sections provide you with foundational concepts that will appear again in later chapters as we delve more deeply into the nature and substance of language.

The origin of language

What a good thing Adam had. When he said a good thing he knew nobody had said it before.

Mark Twain (1935), Notebooks

There is no direct physical evidence for the origin of language in humans and it is not clear how this ability developed. The capacity to speak involved the development of a series of interconnected physical, psychological and social skills and predispositions. Effective communication requires that we are able to exercise fine motor control over our vocal tract, for example, but this skill is shared with a great many mammal species. Yet other aspects of the task of communicating through language, including our ability to anticipate what other people are thinking, are clearly more specialised. Research on the origin of language is now a field shared by specialists with a focus on prehistoric human development; they work in a range of disciplines, including anthropology, archaeology, genetics and psychology.

One of most interesting aspects of language is how languages differ from each other. These differences cut across the full spectrum of what it means to speak a language, from the sounds used by speakers to form words to the different ways politeness is shown, and on to the ways in which important objects are identified and named. Although it is difficult to pin down in objective terms, there is a long tradition of referring to these collective differences between languages as their ‘genius’.

In this book, we decided to select a series of case-study languages that would allow us to explore these differences, while at the same time giving readers the opportunity to become familiar with some languages about which they may previously have known very little. We have chosen languages that vary according to their size and status, and that represent a range of different sound systems, grammatical features, lexical structures and areas of research activity. The languages on which we focus are:

- English (a West Germanic language with a staggering number of speakers included in the almost eight billion people in the world)
- Auslan (the sign language of Australia)
- Gurindji (a Pama-Nyungan language traditionally spoken in the Victoria River district of the Northern Territory, Australia)
- Japanese (an East Asian language and the national language of Japan)
- German (a West Germanic language, with official or co-official status in Germany, Austria, Switzerland, South Tyrol (Italy), Luxembourg and Liechtenstein)
- Mali (Baining) (a Papuan language spoken on the island of New Britain, Papua New Guinea)
- Māori (the traditional language of the original inhabitants of New Zealand)
- Russian (the official language of Russia, Belarus, Kazakhstan and Kyrgyzstan, and the most widely spoken of the Slavic languages, with large speech communities in Ukraine, Moldova, Latvia, Lithuania, Estonia, Uzbekistan and Israel)
- Turung (a Tibeto-Burman language spoken in six villages in Assam in Northeast India).

English is included here because it is usually easiest to learn new linguistic concepts by thinking about them in a language that you already know. So, we often start with English and then dip into these other languages to enable you to experience just how different linguistic categories can be across languages and how differently they can be realised. You can learn more about each of these languages, including the sources of our data, at the end of this chapter and in the online materials for this book.

1.1 Hockett's design features of language

One must let people talk, since fish can't.

Polish proverb

What makes human language special? In the 1960s, the American linguist Charles Hockett developed a list of features that he considered were useful in drawing out the differences between human and animal communication. Many of these are shared with other animals, particularly primates. It is the use of all these features together that makes human language unique. The following list contains a selection of these 'design features' (as Hockett described them), with some comments about their relation to animal communication.

Duality of patterning refers to the fact that language is organised at the level of forms and the level of meanings. Thus, the sounds of English that can be written as <s>, <t> and <a> are forms that can be combined into meaningful sequences, such as as, at, sat, ta, tat and tats (the act of making tatting, a type of lace); in Middle English, there was also the word tas, which meant 'a heap, a pile'. Adding the sound we write as <k> or <c> additionally gives us ask, cat, cats, tack, tacked, tacks, sack, sacked, sacks, task, tasked and tasks. Not bad for just four sounds! (Note that the past tense ending is spelt -ed but pronounced in these words simply as t.) You can read more about the way sounds are distinguished and combined in different languages in Chapters 7 and 8.

Productivity, in Hockett's terms, is the ability to formulate new expressions and utterances based on the words and sentence patterns available in language. In linguistics, productivity also applies to the ability of speakers to form new words in their languages. Since possibilities for the recombination of elements are so rich, the scope for producing new meanings is potentially infinite. In comparison, animal systems of communication are comprised of small sets of calls or sequences of calls with fixed meanings. We explore the creation of words in Chapters 4 and 5 and the rules for creating sentences in Chapter 9.

Even dancing honey bees can't invent new words

The structural and creative aspects of human language are what really set it apart from the communication systems of other animals. As just described, in human language, elements can be combined and recombined in a systematic way to create new forms. It is a very economical aspect of language; that is, discrete units of sound (themselves not meaningful) combine to form meaningful signs (parts of words or complete words), and these then combine and permute to form even larger structures (sentences). The 'moo' of a cow or the 'chutter' of a vervet monkey doesn't show this duality of levels – these signals can't be broken down in that way.

So human languages have a hierarchical structure and, most importantly, are open systems. These discrete parts can be combined and recombined to produce an infinite number of different messages. While some creatures (such as monkeys) have impressive repertoires, they don't seem able to come up with anything new. Their systems are fixed, not open-ended in the way human language is. Even beespeak, it turns out, can't create a word for 'up'!

We humans can talk about literally anything we like, when we want to and where we want to. As Stephen Fry once pointed out, we can be pretty sure that nobody has ever produced the following utterance in the history of human communication: 'Hold the newsreader's nose squarely, waiter, or friendly milk will countermand my trousers.' Humans can even talk about things that don't exist – the man in the moon or the tooth fairy.

Displacement is the ability to talk about things that are not actually present. People spend a lot of their time talking about things displaced in time, in the past or the future, as well as things displaced physically, because they are elsewhere, do not exist (though they are wished for), or no longer exist. Insects that work in colonies – particularly bees and ants of specific types – have been shown to be able to communicate about displaced objects that do exist. Bees can convey information about the location of food (and so show some displacement), but they can't report on that terrific pollen patch they visited last week, or swap stories about great food sources they have known.

Reflexiveness makes reference to the fact that people can use language to talk about language: not only do they know their language, but they can reflect on it. This whole book is a big exercise in linguistic reflexivity, and an important set of tools being presented here is the metalanguage we use – that is, a set of concepts and labels for talking about language. It is not apparent that any other communication systems involve the ability to refer to themselves, or that any other animals indulge in self-reflection relating to how and why they communicate the way they do.

Traditional transmission refers to the fact that children learn language as part of their socialisation into their families and communities. While there is a good deal of debate about the extent to which the ability to acquire and use language is genetically encoded (something we address in Chapter 17), the actual transmission of language among humans is a cultural rather than genetic phenomenon. The communication systems of animals, in contrast, are generally instinctive, there from birth and associated with clear genetic predispositions. Even among animals that show evidence of learning to communicate, there is much less reliance on learning from others than there is among humans. Long-standing research on songbirds, for example, indicates that not only are songs learned along family lines, but in some species song development is responsive to subsequent exposure to different songs. However, birds brought up in isolation will instinctively sing songs (even though they might be a bit odd). This is not the case if a child is brought up in isolation from adults, as we will see when we explore the process of language acquisition in chapter 17.

Can animals lie?

Another of Hockett's design features of language was prevarication, or the ability to lie. There are plenty of examples of animals acting deceptively for the purposes of protecting themselves or their offspring, or to obtain food. One way of differentiating this behaviour from true lying is to assert that lying requires a theory of mind. In relation to lying, theory of mind involves an understanding on the part of the liar that the target of the lie has a mind that can be influenced into a state of knowing something that is false. Then, when the liar acts deceptively, they are doing so with the intention of creating a false belief in the target of their deception.

Based on your experience with pets, do you consider that they could lie? How could you establish their intentions and their beliefs about the state of your mind? To what extent do you think the ability to lie is dependent on language?

We explore the shared assumptions associated with language in interaction in Chapter 14, and return to theory of mind in Chapter 16.

1.2 Signs

But if we had to name anything which is the life of the sign, we should have to say that it was its use.

Ludwig Wittgenstein (1965 [1933–34]), The Blue Book and Brown Books

The duality of patterning described above refers to the relationship between a particular sequence of forms and a specific meaning. This combination of a distinctive set of forms and a specific meaning is called a sign. The study of signs is known as semiotics, a wide field concerned with the many ways in which meaning is expressed and attributed, ranging across the domains of body language, art, music, literature and so on.

The meanings attributed to signs vary according to the culture and context in which they are found. Signs exist when a particular form carries a conventional meaning. These meanings are recognised by others and interpreted in broadly consistent ways. Signs vary in the strength of their relationship to the meaning they represent. In discussing signs, it is useful to distinguish between the signifier, the form being used as the sign, and the signified, the meaning being represented. This distinction was formulated by de Saussure in a series of lectures published in 1916.

Using Figure 1.1 as an example, in the real world, there might be a plant of the type represented by the image. The meaning TREE associated with this image is signified by a range of words (sets of forms that function as signifiers) in different languages: Baum in German, amēnggi in Mali, rākau in Māori and derevo in Russian.

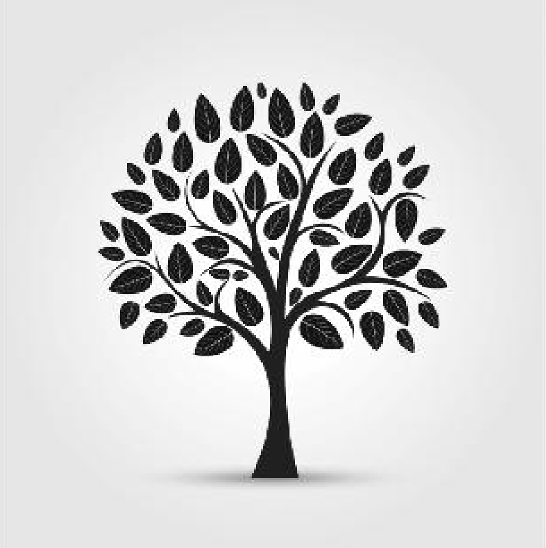


Figure 1.1

Stylised image of a tree

Iconic signs involve a direct representation or imitation of the meaning of the sign. This does not mean that the signifier is necessarily an exact representation, but that some key characteristic(s) of the signified are present in the sign. Our picture of a tree is iconic because it stylistically represents the main parts of a tree for the viewer. Road signs depicting various animals that may run onto the road in different places (kangaroos, deer, moose and so forth) are iconic of the creatures they depict. Iconicity in language includes ideophones (words such as kira kira ‘glitter’, garagara ‘rattle’ and buku buku ‘bubble’ in Japanese), onomatopoeia (roosters say cock-a-doodle-doo in English, kokekokkō in Japanese and kukariku in Russian), and vowel lengthening to emphasise length in storytelling (for example, It went for a looooong time). Many animal postures and gestures also resemble what they are referring to; for instance, when baboons open their mouths, this threat is clearly iconic: the pose looks a lot like biting.

Indexical signs have a direct physical or causal connection to their meaning that can be inferred by the observer. Thus, smoke is indexical of fire: smoke is a signifier and fire is the thing that it signifies. A great deal of modern machinery makes use of indexical signs, beeping and flashing to tell the user something. In English, words such as here and now are indexical of the present time and place. Accents are also indexical, giving information about where someone is from. A lot of animal communication is made up of indexical signs – for example, in ‘beespeak’, the length of the waggle dance correlates with the distance from the nectar.

Symbolic signs are signs in which there is no resemblance between the signifier and the signified. The relation between the signifier and the signified is based solely on convention. The relationship of a yellow background on the animal signs mentioned above and the meaning ‘watch out, one of these creatures might wander onto the road in front of your car’ is a case in point. The fact that in other places animals are placed on white triangular signs with red borders to communicate the same meaning indicates the degree of conventionalisation involved; in other words, people have agreed that this particular sign conveys this meaning. Human communication is highly symbolic; there is no natural and no necessary connection between a word like book and its meaning. It is simply that we are all agreed on calling it ‘a book’ (unless you are Humpty Dumpty in Lewis Carroll’s Through the Looking-Glass, and What Alice Found There – he undermined the very foundation of human language by insisting that words could mean whatever he wished them to mean!).

In linguistics, we are concerned specifically with linguistic signs; that is, the form–meaning relationships found in language. The following sections describe some core characteristics of linguistic signs.

1.3 Arbitrariness of signs

According to de Saussure, linguistic signs are arbitrary. In other words, they are examples of symbolic signs with only conventional relationships between the signifier and the signified. No doubt following de Saussure, Hockett made arbitrariness one of his design features (though it is not unique to human communication). Arbitrariness in language relates to the fact that the same meaning can be expressed in an unlimited way across languages. For example, even onomatopoeic sounds vary across languages – sneezing in Russian is written apchkhi! instead of achoo! This variation is even clearer in non-iconic words. For example, there is nothing apparently iconic or indexical about the relationship between the sounds in the word dog and the meaning it carries. This seems particularly clear when we consider the forms and precise meanings of words for dog in a wider range of languages (see Table 1.1).

Table 1.1 Forms and meanings referring to the concept of dog in a

selection of languages

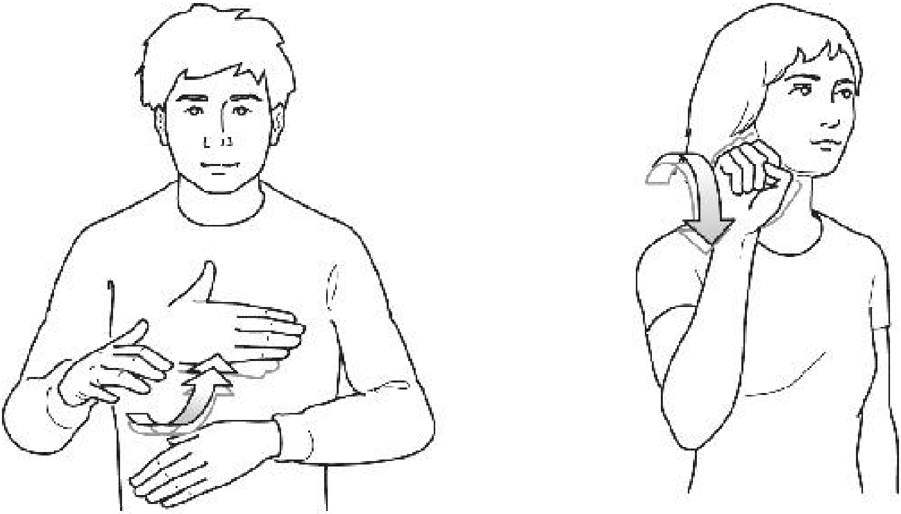
Language	**Form**	**Meaning**
English	dog	Scientific name: canis familiaris (a secondary meaning is to identify male animals of several species including dogs, in contrast to the word bitch for female counterparts in the relevant species)
Māori	kurī	‘dog, quadruped animal (e.g. a deer or kangaroo)’
Gurindji	warlaku	‘dog’
Japanese	inu	‘dog’
German	Hund	‘dog’
Mali (Baining)	vang	‘dog, especially male dog; cf. vaimgi ‘female dog’
Russian	pës	‘male dog’
	sobaka	‘dog (in general); sometimes female dog’

The point de Saussure was making is that there is nothing special about dogs constraining the forms being used. We are forced to conclude, then, that all of them refer to canis familiaris (sometimes among other creatures) by virtue of convention. This is the same for Auslan signs. Even though many of them are iconic, the links between the signs and their meanings are based on customary usage.

Auslan

Within Auslan, there are three different signs for dog. One imitates a dog begging; another, restricted to New South Wales and Queensland, involves patting the thigh as if beckoning a dog; and the third sign appears to imitate the movement of a dog. See the resources listed in the online materials for details.

These Auslan signs for dog suggest actions that are connected with dogs and are therefore not arbitrary in the same way as the spoken words. They are conventionalised as words are in spoken language, but they are also iconic. Some signs of Auslan are arbitrary, but iconicity is a common feature of signed languages generally. For example, the words for cat in Auslan and NS (Nihon Shuwa, Japanese Sign Language) both represent something iconic about cats. As shown in Figure 1.2, the Auslan sign refers to the way people stroke cats, while the NS sign portrays a cat washing itself.



CAT (Auslan) CAT (NS)

Figure 1.2

The signs for cat in Auslan and Nihon Shuwa (Nihon)

Even though we have highlighted arbitrariness as a significant factor in spoken language, there is also a good deal of evidence of iconicity in the way spoken linguistic signs are organised. This is a theme that will recur throughout this book. For example, the English words bash, mash, stash, crash, dash, lash, hash, rash, brash, clash, trash, plash, splash and flash all involve forceful, mostly destructive contact and its results. There is a relation between form and meaning here, so these forms are not simply arbitrary. Word families of these types are always recruiting new forms as they grow. Because of words such as grimy, grotty, grit, grubby, greasy and even gross, English speakers have made a connection between ‘gr’ and muckiness, and they have changed gunge to grunge so that it appears more like the other grotty ‘gr’ words. Individual sounds aren’t supposed to have meanings, but clearly that is not what speakers think. There is a lot more iconicity in our languages than is often recognised!

1.4 Language modalities

Since words are only names for things, it would be more convenient for all men to carry about with them such things as were necessary to express the particular business they are to discourse on.

Jonathan Swift (1726), Gulliver’s Travels, Part III (‘A Voyage to Laputa’)

So far, we have looked at words, which are linguistic signs, mainly in terms of the way they sound. Spoken language is used by the majority of people and is represented in writing and other types of codes. Spoken language makes use of the auditory-vocal channels of communication. In other words, speakers make use of the vocal tract to produce speech noises that are heard by the people they are talking to. As we have just seen, however, human communication doesn’t have to involve vocal signals. Sound just happens to enter into the organisation of human language, but it is not essential. Profoundly deaf people cannot hear – the auditory channel is not available to them; instead, the Deaf can use sign languages, based on the visual-gestural channels of communication. In these systems, signers communicate by making signs using body parts, including the fingers, hands, arms and face, and these are seen (or felt in the case of Deafblind users) and understood by the people with whom they are signing. We make use of Auslan as one of our languages of interest throughout this book, and you can find a fuller account of the language (as well as the tactile form used by Deafblind Australians) in the online materials.

For many years, there was debate about whether sign languages were full languages equivalent to spoken languages in terms of the types of characteristics listed by Hockett and discussed above. There is no longer any doubt about the fact that sign languages are ‘proper’ languages. Moving on from this debate allows researchers to take a really good look at sign languages on their own terms. Furthermore, research into sign languages is revealing, as it allows us to explore human language in a different modality, providing alternative perspectives on a range of issues. For example, the idea of arbitrariness in language seems to be somewhat in conflict with the greater degree of iconicity in sign languages. The important point about signs (whether spoken or gestural) is not so much whether they are iconic, but the fact that they are all conventionalised. This observation also supports our earlier account of sets of words (such as -ash and gr-) that are based on sound similarities. These sorts of word families are iconic, but they also invoke a conventionalised, and hence an ‘agreed-upon’, association between sound and meaning. In short, language works because we are all agreed on what signs mean (whether they are iconic or arbitrary), and this is the characteristic shared by all linguistic signs of either modality.

It is important to clarify that sign languages are different communication strategies from the gestures people use as they speak. Co-speech gestures are spontaneous, non-conventionalised supports to communication rather than complete language systems in themselves. Similarly, acting out a story by miming is completely different from telling a story using sign language. Miming and gesture can be very helpful when there is no common language, but these activities fall far short of constituting a sign language in terms of the criteria we have outlined here.

Writing is a way of representing an originally spoken language in a visual mode. It would have started life as a simple memory jog for speakers (inventories, records and so on). These days, as English speakers, we are so immersed in writing that it is hard to imagine a time when the written language was nothing more than an optional extra. However, we have to consider writing as secondary to speech, as well as to sign language, for reasons including the following.

- Writing systems were invented after humankind began using speech.
- Spoken language is found in all human communities; writing is not.
- Children learn to speak before they learn to write.
- Spoken language is more frequently used than written language, even in literate communities.
- Spoken language is more expressive than written language.

There are a number of different ways to write the words of a language.

1 Logographic systems have symbols for each word; for example, the ‘characters’ of languages, such as Chinese and the Japanese kanji. English has many examples of logographs: 3 is a logograph for ‘three’ and $ is a logograph for ‘dollar’.

2 Syllabaries, such as Cherokee and Japanese (specifically, the kana), have symbols for each syllable. So syllabic systems tell us how to pronounce the words, but logographic systems don’t (an example in English might be something like bar-b-q).

3 Alphabetic systems have symbols that represent the separate sounds that make up a syllable; the earliest known system approximating an alphabet was North Semitic (developed around 1700 bce in Palestine and Syria).

In Chapters 7 and 8 we return to the topic of how sounds are represented in writing. Here we are concerned more with the ways in which writing differs from speech as a means of communication. There have been many attempts at identifying key differences between speech and writing as they are typically used. Table 1.2 shows a few widely noted differences.

Table 1.2 Differences between speech and writing as they are typically used

	Speech	**Writing**
Position of audience	**Present**	**Absent**
Feedback from audience	**Available**	**Not readily available**
Use of paralinguistic cues (loudness of voice, rate of speech, facial expressions and so on)	**Can be used**	Cannot be used – word choice, punctuation, emoticons and emoji are alternative options
Permanence	**Usually transitory**	**Often long-lasting**
Context-dependence	**Yes**	**No**
Association with literary genre	**Less strong**	**Strong**
Use of standard variety of the language	**Less expected**	More expected (especially in formal writing)
Evidence of editorial efforts	**Little attention**	**Extensive attention**
Planning time	**Real time**	**Often extensive**

The discussion in Chapters 13 and 14 about different ways of studying texts and conversations draws out some of these differences further. An interesting current area of study relating to this general topic is the use of language on the internet. Facebook chat is an obvious example of a hybrid form of communication; language is communicated in written forms, but the interactional patterns involved have many features in common with face-to-face conversations. Texting on mobile phones and the use of emoticons and emojis in particular (see Figure 1.3) also reflect developments that blur the traditional distinctions between speech and writing.

Figure 1.3

Pack of emoticons

1.5 ‘Infinite use of finite means’: Rules and paradigms

We have now established that meaning in language is conveyed by signs, which at the basic level correspond to words. But we don’t normally speak in single words; rather, we organise words into larger sets and use these to communicate more complex meanings. This is the topic of Chapter 9. For now, we want to give you an overview of how speakers manage in this process.

Let’s start by thinking about how words are organised into strings. It is pretty obvious when you are reading words that they are organised in a linear way – one word follows another across the page. Speech is linear too, with words being uttered one after another in time. The order of the words is organised by special sets of rules, and these can vary across languages. For example, in English sentences the ordering is strict:

(1.1)

My daughter’s friend caught a very big fish.

In this sentence, My daughter’s friend is the subject (the ‘catcher’) and it must precede the verb caught; a very big fish is the object (the ‘catchee’) and it must follow the verb. Furthermore, the order of the elements that modify nouns are also fixed. In the phrase a very big fish, the noun fish occurs last, and the preceding elements are used to tell the reader or hearer additional information about the fish. As example 1.2 shows, other orders are not acceptable in English (and the convention is to indicate this with an asterisk *):

(1.2)

a big very fish

*big very a fish

*very a big fish

*a very fish big

*a fish very big

*fish a very big

The above two examples show how English speakers follow rules in organising words into strings. As we will show in more detail in Chapter 9, there is quite a bit of variation across different languages regarding what kinds of principles are at work. For example, the rules in Mali are slightly different to English. Although the order of the subject, verb and object is the same, the modifiers in the noun phrase are not. In example 1.3, the word morki ‘big’ follows the word thinēmki ‘fish’ and also requires a linker, ama, translated here as ‘it be’.

(1.3)
[illegible]
[illegible]
‘My daughter’s friend caught a very big fish.’

(Literally: ‘a large fish that was big (for a large type of fish)’

Words are also organised in relation to each other in terms of their meanings. A favourite pastime of descriptive linguists is to construct paradigms. These are sets of different forms of the same word, often arranged as a table to show how their meanings are related to each other. Table 1.3 shows a pronoun paradigm for English.

Table 1.3 Pronoun paradigm for English personal pronouns

A	**B**	**C**
Subject	**Object**	**‘Self’**
I	me	myself
we	us	ourselves
you	you	yourself/yourselves
he	him	himself
she	her	herself
it	it	itself
they	them	themselves

Notice how any word from the ‘Subject’ pronoun column can go into slot A in example 1.4, and any word from the ‘Object’ pronoun column can go into slot B. However, if the identity of the subject and the object were the same (in other words, the subject did something to themselves), then a pronoun from column C and the same row as the ‘Subject’ pronoun would have to be used in slot B instead.

(1.4) [A] locked [B] out of the house accidentally last week.

Paradigms do seem to be psychologically real for people, and for that reason they can be a handy teaching device. Those of you who have learned a foreign language have probably encountered sets of words like those in Table 1.3. As we will see in Chapter 11, there are also certain sorts of changes that take place within paradigms. So, linguists’ preoccupation with patterns and paradigms is well founded!

We have now established that linguistic signs are organised into larger units (phrases and sentences) that express more complex meanings using rules. Anyone who is a speaker of a language will have a pretty good sense of these rules and exactly what they allow. Most readers of this book will be able to reliably distinguish between the three types of English sentences:

1 grammatical sentences – sentences that are easy to interpret and intuitively correct

2 ungrammatical sentences – sentences that are not easy to interpret and are intuitively wrong

3 semantically troubling sentences – sentences that are grammatical but that do not make sense (so they are ‘wrong’ in terms of conflicting with our expectations about how the world works, although they do follow the rules of the grammar).

See whether you can identify the status of each of the following sentences.

(1.5)

(a) Marnie coughed because it was so dusty.

(b) The chair collapsed because it was so dusty.

(c) Chair collapse be very dusty.

(d) Chair-collapsing is hazardous to your health.

(e) Flowers like hazardous to your health.

(f) Mario is a champion internet browser.

Teaching animals human communication

In the past 50 years, there have been many attempts at teaching apes human language. Those in the 1930s and 1940s were a failure, probably because the animals were made to use their vocal organs; the ape's vocal apparatus is structured quite differently and is fundamentally unsuited to the production of vowels and consonants. In the 1960s, a series of famous projects used computer consoles, plastic shapes on boards or a type of sign language. According to their trainers, these creatures learned hundreds of words and could string them together in meaningful ways. The most famous studies involved Koko (a gorilla) and Nim Chimpsky (the chimp's name is clearly a dig at linguist Noam Chomsky, who was well known at the time for holding the view that humans alone were 'hard-wired' to develop language). Both apes were taught a kind of modified American Sign Language.

A little later, an even more interesting primate, Kanzi (a baby bonobo, or pigmy chimp), became well known. He accompanied his mother to sessions where she was taught language through symbols on a keyboard. Although Kanzi appeared to show no interest in these lessons, he spontaneously started to use the symbols and soon outstripped his mother in ability. His linguistic skills are reputed to have been impressive – around 250 symbols, which he seemed to be able to put into simple grammatical sentences. He also consistently produced distinct noises when referring to things such as bananas and grapes. Kanzi was the first of these ‘talking’ primates to be immersed in language in a more natural setting and from a very early age; in other words, he didn’t learn through direct training, but simply through exposure – just as human children do.

1.6 Prescriptive and descriptive approaches

A Grammar book does not attempt to teach people how they ought to speak, but on the contrary, unless it is a very bad or a very old work, it merely states how, as a matter of fact, certain people speak at the time that it is written.

H.C. Wyld (1925), Elementary Lessons in English Grammar

The interesting thing about rules and paradigms is that they vary from person to person, and particularly from one group of people to another. Although English is spoken in many countries, the form of English spoken in each place is to some extent unique. For example, in Ireland one can say all of the following.

(1.6)

(a) You’re after ruinin’ me.

(b) I have it forgot.

(c) I think the younger generations are gone idle over it.

(d) I’m not in this place long …

How many of these are acceptable in your variety of English? If you heard someone speaking in this way, what would you think? It is likely that unless you also speak Irish English, you’ll find at least some of these examples ungrammatical. It is only by recognising that there are distinct varieties of English with different rules that we can avoid writing off these differences as mistakes. In order to distinguish between approaches to language that impose a standard and approaches that focus on observing what is happening, linguists have drawn a distinction between prescriptive and descriptive approaches.

Prescriptive approaches place value judgements about language at the centre of the frame. Often associated with standard varieties of major languages, prescriptive approaches involve the documentation and promotion of norms about language use, and (often) the denigration of behaviours and practices that fall outside these norms. In this approach, there is no room for reflection about the values attached to different varieties. In fact, there are only two varieties: the good one and the bad one!

Descriptive approaches, on the other hand, focus on describing what people say and do with language, and on trying to learn about language more generally, based on these observations. Most linguists try to take this approach, which also involves avoiding personal value judgements about how language is being used. In this approach, we may observe the value judgements of others – this is an important aspect of sociolinguistics.

Given the importance of standard languages in a person's education, work and social life, there is a place for awareness about language norms and the values attached to them. Although linguists are generally very comfortable with the descriptive approach, and enjoy the weird and wonderful ways in which languages like English can be used, they are also aware of the disadvantages that accrue when English is not used appropriately in contexts with strict conventions. Sociolinguistics focuses on this kind of interaction between language, and the structure and functioning of society; in Parts 3 and 4 of this book, we explore important sociolinguistic aspects, such as prestige and solidarity.

Finally, it is interesting to note how both prescriptive and descriptive approaches can position speakers in less powerful roles than the expert linguist. It is useful to be aware of this problem if our work as linguists is to have a positive impact overall. In a prescriptive approach, individuals are expected to submit to more widely held norms, whereas in the descriptive approach the question of compliance does not arise. Nevertheless, the linguist as the expert still holds a good deal of power in saying what language is really like.

An interesting case study that shows the difficulties of maintaining a straightforward distinction between descriptive and prescriptive approaches to language is the Oxford English Dictionary. From a descriptive perspective, the dictionary is:

- a big collection of true facts about words – an exhaustive description of English vocabulary
- constructed empirically

- sometimes used as a model but not intended to be prescriptive
- without political significance.

According to this view, word usage that diverges from that reported in the dictionary must be either mistaken or an example of language change not yet incorporated into the dictionary. However, from a prescriptive perspective, the dictionary is often read as being:

- a big collection of normative assertions about words
- a final court of appeal for questions about English
- the best model available
- an important symbol of English language prestige.

In this view, in addition to making mistakes, people might deliberately flout the conventions represented within the dictionary. In a funny way, the prescriptive view of language allows speakers and writers additional agency that is not available in a straight descriptive view, since rules must be established before they can be broken.

Should a dictionary contain only the ‘good’ words?

In the November 1857 meeting of the Philological Society in London, Richard Chenevix Trench delivered what was to become the most important paper in the history of English lexicography. Due to his dissatisfaction with current dictionaries, Trench called for two reforms. He wanted a dictionary that captured the actual lifespan of words, with quotations that would shed light on their journeys – a task that would have seemed whopping at a time well before scanning devices and the internet! Most importantly, he had a vision of a dictionary that was an inventory of language, one that captured all words (old, new, good and bad) and gave equal treatment to all usages. Here is an extract from this famous paper.

A Dictionary then … is an inventory of the language … It is no task of the maker of it to select the good words of a language … If he fancies that it is so, and begins to pick and choose, to leave this and to take that, he will at once go astray. The business which he has undertaken is to collect and arrange all the words, whether good or bad, whether they do or do not commend themselves … He is a historian of it [the language] not a critic.

Trench was expressing the idea of the dictionary as a non-prescriptive historical record of the language and it is this practice that now characterises English lexicography. To drop or exclude a form considered sub-standard provides an inaccurate representation of the language and therefore diminishes the value of the work – you might draw a comparison with a zoologist who studies and records only the physiology, anatomy and behaviour of animals of which the zoologist approves! Mind you, words have to have some longevity before they are placed in a dictionary, and this is the challenge for modern lexicographers – deciding whether or not an expression will endure.

1.7 The ideal and the real

The picture of language that we have been building up in this chapter so far assumes that language is an abstract system of signs that we combine by following established rules. This is a powerful way of looking at language. It allows us to write grammars and dictionaries, and to know many things about how language works. For some researchers, these patterns are understood to directly reflect the way language is organised in the brain. However, there is lots of evidence that the systems of language modelled through these approaches do not always match up with the things that people actually do when they talk, sign or write to each other.

We have just observed that language practices that might be seen as mistakes from a prescriptive perspective may simply represent differences in usage in a descriptive framework. For researchers who make a close association between descriptions of language systems and mental representations of language, it is also necessary to draw a distinction between the ideal and the real, or the things people know as opposed to what they actually do. Noam Chomsky drew our attention to these differences, labelling them competence and performance.

In this approach, language acquisition is viewed (among other things) as involving the development of grammatical competence over time. Once it is established, grammatical competence is considered to be a fairly stable thing. Only adverse events, such as brain injuries, are likely to undermine grammatical competence once it has been established. Competence relates to what speakers ‘know’ about how language is meant to work. It has to do with the ‘rules’ we have learned and internalised unconsciously by simply being part of a speech community; even though we might not be able to say what these are, we know when something goes wrong. For example, our ability to identify one of the following example sentences as ungrammatical is based on our grammatical competence in English.

(1.7)

(a) Marnie coughed because it was so dusty.

(b) The chair collapsed because it was so dusty.

(c) Chair collapse be very dusty.

Performance, on the other hand, is susceptible to all sorts of contextual influences. If we are tired, distracted, upset, nervous and so on, this can have an effect on whether or not the language that we produce matches what we might consider to be correct or competent use of language – everyone's 'tingue slops' on occasion! In Chapter 16, we will look at a range of speech errors and consider how these errors allow us to model the production of language in the brain.

Linguists also talk about communicative competence. This is the ability to use language appropriately according to the communicative norms of our society. The ability to apologise, tell a joke or invite someone to dinner are examples of interactions that involve communicative as well as grammatical knowledge and skills. Again, our ability to do these things well depends on a range of contextual factors, together with our basic understanding of what is appropriate at any given time. Tiredness, distraction and strong emotions (or drink!) are all factors that can interfere with our ability to communicate effectively.

1.8 Language in its social context

One of the things Ford Prefect had always found hardest to understand about humans was their habit of continually stating and repeating the very very obvious, as in It's a nice day, or You're very tall … At first Ford had formed a theory to account for this strange behaviour. If human beings don't keep exercising their lips, he thought, their mouths probably seize up. After a few months' consideration and observation he abandoned this theory in favour of a new one. If they don't keep exercising their lips, he thought, their brains start working.

Douglas Adams (1979), The Hitchhiker's Guide to the Galaxy

The idea that language is systematic and rule-governed can get us a long way in describing how language works regarding form and grammar, but clearly there are many aspects of the human experience of using language that remain to be explored.

The character described in the above quotation is Ford Prefect, the hilarious creation of Douglas Adams. Ford comes from a small planet somewhere in the vicinity of Betelgeuse but has visited Earth in order to carry out research for an article he is writing for The Hitchhiker's Guide to the Galaxy. As Ford so astutely observes, many linguistic exchanges regularly take place without involving any transfer of information at all. A lot of the time, people don't seem to be communicating anything particularly and much of their chit-chat involves stating the obvious – all that talk about health and the weather, and the linguistic formulae that fill up conversations (the pleases, the thank yous, the you're welcomes, the pleased to meet yous, the bless yous and so on). They don't advance the conversation. So what are they doing there?

The kind of social chit-chat (or small talk) to which Ford Prefect is alluding is used to establish social rapport during an encounter. Most human interaction is non-hostile. Usually we are polite whatever we might be feeling deep down, and these sorts of ritual linguistic exchanges are all part of this. They are a way of signalling that there is no animosity. Although not all cultures use language in this way (as we'll see in Chapter 14), these sorts of verbal routines are often used to express important information to do with an individual's social characteristics, relationships and values.

Clearly, the relationships between language users are important, and a rule-bound model of language only gets us so far. How does the speaker communicate effectively with the hearer in a complex social world? What types of reasoning does the hearer use to make sense of the speaker's words? What is there in the context, besides the words themselves, to guide the hearer? There is so much to communication and social contact that falls outside of the concepts of signs and rules, competence and performance. These concepts do not provide us with the means to describe why swear words are so upsetting to some, nor can they explain why some language varieties are valued more highly than others. Processes of language change and the motivations that drive these are also difficult to describe in terms of signs and rules alone.

Other ways of looking at language place the act of communication at the centre of the picture and explore how people use language for specific ends. The significance of placing people in the scene cannot be over-emphasised, since people, with all their prejudices, preoccupations, fears and desires, significantly complicate the picture. People exist in space and time; their plans and choices are defined by their history and their context. Their intentions are not always clear – even to themselves. They exaggerate, lie and make creative use of metaphors. It is clear that understanding the ways in which they behave and, specifically, the language choices that they make requires a much more open and socially oriented approach.

1.9 Language across time

In studying language, some linguists focus on the state of a language at a particular period (a snapshot in time), while others take a historical perspective and analyse the changes that have taken place across time (a series of snapshots over time). This contrast in approaches is also referred to as a focus on either synchronic or diachronic views of a language. Both are important for a complete picture of any language.

Languages change over time and the traces of older versions of a language can resonate in the present. So, even when the focus is on the current state of affairs in a language, it is often useful to consider the historical context in order to understand the way things work. For example, English is a Germanic language, but it has an enormous amount of French vocabulary. Knowing the history of the Norman invasion of England helps us to understand why this is so. When William the Conqueror defeated King Harold at the Battle of Hastings in 1066, he subjugated the rest of south-east England. Within ten years, all of England was under William's control – politically and linguistically, this was a French conquest of English. To illustrate, consider the following simple recipe for haddock in a tasty sauce. We have highlighted the French words.

Haddocke in Cyuee. Shall be yopened & ywasshe clene & ysode & yrosted on a gridel; grind peper & saffron, bred and ale mynce oynons, fri hem in ale, and do therto, and salt: boille hit, do thyn haddok in plateres, and the ciuey aboue, and ghif forth.

Haddock in civet [= stew]. (It) should be opened and washed clean and boiled and roasted on a grill. Grind pepper and saffron, bread and ale, mince onions, fry them in oil, and combine and salt: boil it, put the haddock on plates and the civet on top, and serve.

There are 48 words in this recipe and thirteen of them are French in origin. The everyday vocabulary remains Germanic; this includes grammatical words, such as conjunctions (and), auxiliaries (shall, be), pronouns (hit, hem), prepositions (in, on), basic ingredients (bread, ale, salt, pepper) and basic preparation verbs (wash, open, do, give).

An eye on history helps explain pairs of words borrowed twice at different periods. Some of these words are listed below. Note how many of these forms also changed in French.

hostel (c.1250) < Old French hostel, hostel	hotel (c.1765) < French hôtel
liquor (c.1225) < Old French licur, licour, likeur	liqueur (c.1742) < French liqueur
ticket (c.1528) < Old French estiquette	etiquette (c.1750) < French étiquette
troop (c.1545) < Old French trope	troupe (c.1825) < French troupe

We focus on the historical perspective in detail in Part 3 of this book and look at the interaction of language change with social conditions and cultural developments.

1.10 The case-study languages

Reason is common, the spirit in each language has its own particular form.

Jean-Jacques Rousseau in L’Émile (trans. by Pountain, 2008)

In the following sections, we provide a brief account of each of the case-study languages on which we draw to illustrate the many different linguistic concepts discussed in this book. You can find more information in the online materials.

Auslan

Insofar as it may contribute to the redefinition of what is ‘language’ or what is properly ‘linguistic’, the short history of the study of signed languages belies its relative importance to linguistics.

Trevor Johnson and Adam Shembri (2007), Australian Sign Language

There is a lot of misunderstanding about sign languages like Auslan. Many people have the idea that they are invented languages. Certainly, there are auxiliary signing systems that have been devised by linguists and educators (and recent times have seen the creation of such versions as signed English and finger spelling), but real sign languages are not invented. They evolve naturally within communities of individuals who are Deaf and, with time, the children of these communities might acquire the sign as a first language.

There are many different sign languages around the world – at least 103 known languages (and probably many more that haven't yet been documented). Some of them are related, and these may or may not be mutually intelligible. The sign language of Australia, Auslan, is closely related to New Zealand Sign Language (NZSL). Both developed from varieties of British Sign Language (BSL) that Deaf migrants and teachers of the Deaf from Britain brought with them, and these connections were strongly maintained from settlement up until the middle of the twentieth century (Figure 1.4).

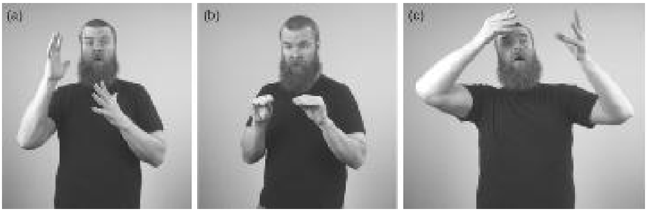


Figure 1.4

Auslan signs for AUSLAN, TEACH and LEARN

Spoken languages are made up of a range of distinctive sounds, whereas sign languages are organised around different forms of signs. Sign language researchers have identified the handshape, location, orientation and movement of the hand as key aspects. These multiple and dynamic aspects of signs are difficult to represent in writing. For this reason, pictures, sometimes containing arrows to indicate the direction of movement, and video clips are the best ways to understand the form of signs. In order to talk about how signs are combined, they are often represented by words in capital letters. For example, compare the signs for CHAIR and COUCH in the Auslan Signbank Dictionary online.

English

English? Who needs that? I'm never going to England.

Homer Simpson in The Way We Was

English is a West Germanic language (of the Indo-European language family), which is a heritage it shares with Dutch, Frisian and German. We discuss the relationships in more detail in Chapter 11. Figures for English speakers these days are hard to determine. The world's population is now almost eight billion and a staggering proportion of these people are regular users of English. It is estimated that there are more than 379 million people who speak English as their first language, and upward of 750 million who speak it as a second language. There are also millions of consumers of English, who are increasingly using it as their international language, both in work and at play. However, the numbers here are hard to estimate because we are potentially dealing with such a wide range of abilities, from fully competent speakers to those millions of semi-speakers with just a smattering of English. Given the great diversity of Englishes used across the world, it is difficult to describe what 'English' is like. We can say that the language has around 44 distinctive sounds, but of course the variation is huge – even between the so-called native varieties.

Over the years, English has become increasingly analytic in its grammar, which means that it relies on features such as word order and free-standing forms to convey grammatical information. Only a handful of the endings (inflections) remain; currently, an English noun like duck can have two forms (duck-ducks), verbs generally come in fours (quack-quacks-quacked-quacking) and some adjectives come in threes (fluffy-fluffier-fluffiest). However, as the inflectional system continues to simplify, separate words are becoming increasingly important (fluffier and fluffiest versus more fluffy and most fluffy).

The English lexicon is thoroughly mongrel in nature (with expressions from more than 120 languages). Another defining feature is the layering effect of stylistic levels. This is the result of contact with other languages: ask (native English) versus question (French) versus inquire (Latin). There are hundreds of examples of this kind of stylistic patterning, ranging from the colloquial through to the formal, from everyday to highly technical, and from the general to the specialised.

The work of Crystal (2004) is a very readable overview of the history of English and books on English as a global language, and English grammar and its usage are in abundant supply.

German

German is spoken by approximately 100 million people in more than fifteen countries around the world. It has official status in Germany, Austria, Switzerland, Liechtenstein and Luxembourg; it also has regional status in parts of Denmark, Belgium, Italy and Romania, and substantial communities of speakers in France, Hungary, Russia, Poland and even in North America and Namibia.

Like English, German is a West Germanic language. There is a standard variety of German (so-called High German or Hochdeutsch) that evolved as a written language from the time of Martin Luther. (Note that the ‘high’ in High German is a geographical reference to the group of dialects spoken in south and central Germany between the Alpine regions and the north German plains, contrasting with the very different Low German varieties spoken on the plains.) The written variety of German is promoted in schools and used in law courts and government institutions. However, there are different national varieties; the standard differs in Austria, Switzerland and, as some argue, even within Germany itself.

Local dialects continue to be strongly used in face-to-face interaction and regional differences can be striking. Near the German and Dutch border, German speakers using their own ‘home-grown’ varieties can communicate better with neighbouring Dutch speakers than German speakers in the south of Germany, who speak southern varieties of German.

Compared to its relative, English, German has retained many more of its inflections (and one of the puzzles is why the siblings have been evolving at different rates). German still has a grammatical feature of nouns known as gender – all German nouns fall into one of three classes: masculine, feminine and neuter, and this also affects the form of those words that modify the nouns. It also has case, which involves grammatical markers indicating the relationship of nouns to each other and to the verb; in other words, they signal who is doing what and to whom. German word order is less fixed than English, and one striking feature is that verb forms are sometimes found at the end of sentences.

In the text, we illustrate a number of linguistic concepts with examples, examining Pennsylvania German, the variety now spoken in Pennsylvania in the United States and Ontario, Canada, by descendants of late seventeenth- and eighteenth-century immigrants from parts of Germany and Switzerland. Kate Burridge has been working with Pennsylvania German since 1986 (Figure 1.5). Here we would like to acknowledge our appreciation to this speech community for their knowledge, for their continued friendship, and for their time and their patience in answering all Kate's pesky questions!

Figure 1.5

Members of the Pennsylvania German community

Gurindji

Languages differ in what they force us to say, in the kind of information which their grammatical structure obligatorily carries.

Yaguello (1988) (trans. by Poutain, 2008)

Gurindji is a Pama-Nyungan language traditionally spoken in the Victoria River district of the Northern Territory, Australia. The Gurindji people held a prominent place in Australian politics in the 1960s and 1970s following the Wave Hill walk-off, as they struggled to assert their rights to decent wages and to ownership of their traditional lands. They now predominantly live in the communities of Kalkaringi and Daguragu in the Northern Territory. The number of speakers is small. Australia's languages are disappearing at a distressing rate, and Gurindji is among the most highly endangered. Of the original 250 languages spoken in Australia, only around 100 languages remain today and most, like Gurindji, are in a precari-ous position, with small numbers of only older speakers.

In Gurindji, grammatical information is always conveyed by suffixes; the language has no prefixes. As is typical of Australian languages generally, word order in Gurindji is flexible and is determined by discourse considerations such as: what part of the message speakers want to highlight, what they think is most important, what they assume their audience already knows, and so on. Gurindji has an extensive spatial vocabulary expressing both the compass direction or orientation and relationships to rivers and other features of the landscape.

Even though the language is no longer widely spoken, Gurindji 'ways of knowing' are continuing through a new language that has been developing in the Gurindji community since the 1970s, known as 'Gurindji Kriol'. This is a mixed language, combining elements of the traditional Gurindji language with Kriol, an English-based creole language that is widely used in Aboriginal communities in the Top End.

Figure 1.6

Lipananyku, the site of the first cattle station on Gurindji country

The Gurindji speech community illustrates the kind of the complex multilingualism typical of Aboriginal communities in rural Australia: traditional languages (such as Gurindji) continue to be spoken by some people, alongside other languages (like Kriol, Aboriginal English and Standard Australian English), as well as new mixed languages (such as Gurindji Kriol) that have arisen though the interaction of all these languages. Many speakers are skilled at moving along a continuum, altering their speech to suit the situation and audience: at one end the mixed languages (with forms and structures of a number of languages), at the other an English variety fairly typical of rural Australia, and in between varieties of the English-based creole.

Research has been conducted with members of the Gurindji community since the 1970s by a number of linguists and literacy workers through the Australian Institute of Aboriginal and Torres Strait Islander Studies (AIATSIS), the Summer Institute of Linguistics (SIL) and the Diwurruwurru-jaru Aboriginal Corporation (the Katherine Regional Language Centre). Our thanks to Felicity Meakins, a researcher with long-standing links to the community, for facilitating communications on our behalf.

Japanese

Even a brief survey of grammatical issues leaves one somewhat in awe at the extraordinary variety of patterns that exist in the languages of the world.

David Crystal (2010), The Cambridge Encyclopedia of Language

Japanese has around 132 million speakers, some 127 million of whom reside on the islands of Japan (where it is the national language), and the rest in other parts of the world (notably the United States and Brazil). There is a standard variety of the language based on the dialect spoken in the higher-class areas of Tokyo. There are many other varieties (especially on the Ryūkyū Islands), the grammars of which diverge significantly from the standard, to the extent that they are not even mutually understandable.

Japanese is distinguished by its different writing systems, often used in combination with each other. One has symbols for individual words; they are the so-called ‘characters’ of the kanji (based on the Chinese characters). Another system has symbols for each syllable; this is the katakana script (for foreign language words) and the hiragana script (for native words). Japanese also uses symbols to represent the separate sounds; this is the Roman alphabet with which English speakers are most familiar, introduced by missionaries from Portugal and Spain in the sixteenth century. For simplicity, we are using this last orthography, so that those of you unfamiliar with Japanese can approximate the pronunciation.

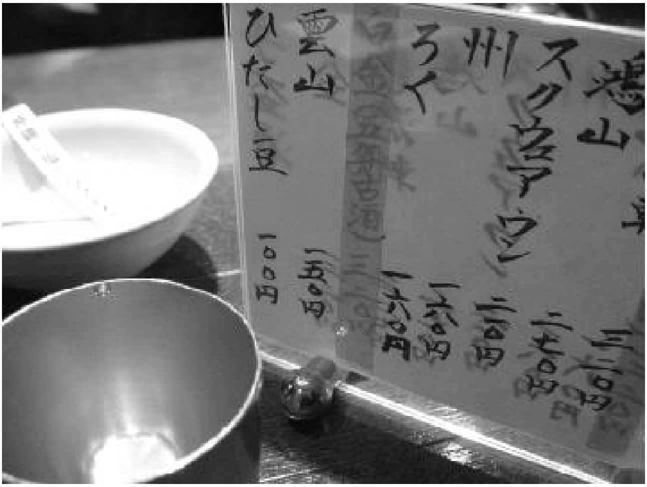


Figure 1.7

Sake menu

Like Gurindji, Japanese is called an agglutinative language. It makes extensive use of endings, in this case on verbs and adjectives, which inflect for tense (for example, taberu '(I) eat' = tabe eat + -ru non-past). Japanese verbs always go to the end of the clause, with subjects usually positioned before objects (although objects can appear earlier). The language has been described as topic-prominent, which means there is a strong tendency to state the topic of the sentence up front, separately from the subject.

Japanese is most famous for its complex system of honorifics, where verb forms, pronouns and vocabulary indicate the relative status of the speaker, the listener and any other persons mentioned. All languages have politeness conventions of some sort, but in Japanese they saturate every level of the language, particularly the grammar.

The influence of Chinese culture on Japanese is clear in the very high proportion of Chinese loanwords in the language. In fact, around 60 per cent of the vocabulary is borrowed from Chinese or constructed according to Chinese patterns (kango words). Similar to the stylistic levels found in English, these words are more formal or more academic than the native Japanese words (yamato words). These days, English is the biggest contributor to the Japanese lexicon (for example, rasshu awaa 'rush hour').

Mali (Baining)

Ut ma Malichēna da vandi amēr gēlēm ut nēvēt autha methomon da sok uthi sangar mamēr pētha ngēt ma tamon nanēgt mermer. (We Mali people must be proud of our language and we must really hold on to it and speak it forever.)

Julius Tayul (in Tayul, Stebbins & Marunga Mali Language Project, 2004), Asēk dēchama mambu nēva bang (Stories and Songs from the Village)

The interior of the Gazelle Peninsula on the island of New Britain in Papua New Guinea is occupied by the Baining ethnic group. This group speaks a set of related Papuan languages, known collectively as the Baining language family. The languages are Kaket, Kairak, Ura, Simbali and Mali. We focus on Mali, which is spoken in the south-eastern quadrant of the Peninsula and has around 2200 speakers.

Linguistic diversity in Papua New Guinea is under serious threat and, like many other languages in the world, Mali is currently being pressured by larger languages of wider communication, particularly Tok Pisin, an English-based creole that is the lingua franca (or common language) of much of the country.

One particularly interesting feature of Mali grammar is that the majority of verbs show changes to the initial sound to reflect tense marking – for example, a verb meaning ‘call, read’ has the forms tes (present) mes (past) and thes (future). (Here you might compare those English verbs that change their vowel to indicate past time (sing-sang-sung.) The Baining languages are also unusual because they have an intricate noun class system in which a single root can be used as the basis for a range of words: amēng ‘wood, tree’, amēng-ka ‘slender tree’, amēng-ki ‘large full grown tree’, amēng-ini ‘stick’, amēng-ēm ‘tree stump’, amēng-igl ‘plank’, amēng-vēt ‘pole’, amēng-ia ‘large log’.

Mali vocabulary shows the effects of borrowing from a number of languages, especially Tok Pisin, English and Latin. In the case of the last two, the expressions probably entered the language via missionaries in the Catholic Church; however, because Tok Pisin is an English-based creole, it is sometimes impossible to tell whether expressions (such as bechi ‘bag’, sensocha ‘chainsaw’, Deo ‘God’, salve ‘greetings’) came into Mali directly from these languages or via the creole route.

Research on Mali has been conducted by Tonya Stebbins. When she first met people from the Mali community and asked what people would like to gain from the research, Tonya was consistently told that they wanted other people in the world to know about them and their language. We include Mali here both because it is an opportunity to follow through on this request and because the language has a range of wonderful features that are interesting to explore. We would particularly like to acknowledge the central contribution that Julius Tayul made in preparing the materials on Mali presented here.

Māori

Ko taku reo taku ohooho, ko taku reo taku mapihi mauria. (My language is my awakening, my language is the window to my soul.)

Māori proverb

Māori is the traditional language ‘of the original inhabitants of New Zealand. It is closely related to Tahitian, and more distantly related to Hawaiʻian and Rapanui of Easter Island.

The language has been in decline for many decades as the result of a long-standing policy of assimilation directed towards the Māori population by European missionaries and settlers, and their descendants. However, since the 1970s there have been increasing efforts to revitalise and extend the use of Māori in New Zealand. The establishment of the Te Kohanga Reo (‘language nests’ for children) and Kura Kaupapa Māori (Māori language medium schools) have had promising results for the people involved, and they are being used as models for revival or revitalisation movements elsewhere in the world.

An important venue for the use of Māori is the marae, or meeting place. This word refers to an established meeting place in which the community can gather. Traditional literary forms associated with these places include: the karanga, a call of welcome by women to an approaching group; the haka, a set of calls and movements performed by men; and the whaikōrero, or formal speeches. There is also a tradition of singing that continues to evolve today.

Vowel length is distinctive, so the meaning of the word changes depending on whether the vowels are long or short. Compare kēkē ‘armpit’, kekē ‘to creak’, and the loanword keke (cake). Parts of speech in Māori can be difficult to distinguish, but typically are said to include nouns, verbs, adjectives (often included in the group of so-called stative verbs that describe a condition of being rather than an action or an experience), location words, pronouns and proper names. Typical of other Polynesian languages, Māori has a wide range of grammatical particles; for example, a number of verbal particles indicate information to do with the timing of the event to which they relate (for example, ka indicates future time in the following sentence: Ka haere au, literally, ‘future go I’). As this sentence shows, another distinctive feature of Māori grammar is that verbs precede subjects, and modifying words follow the noun; for example, pukapuka whakamere ‘book interesting’ (compare this with English where modifying words precede nouns).

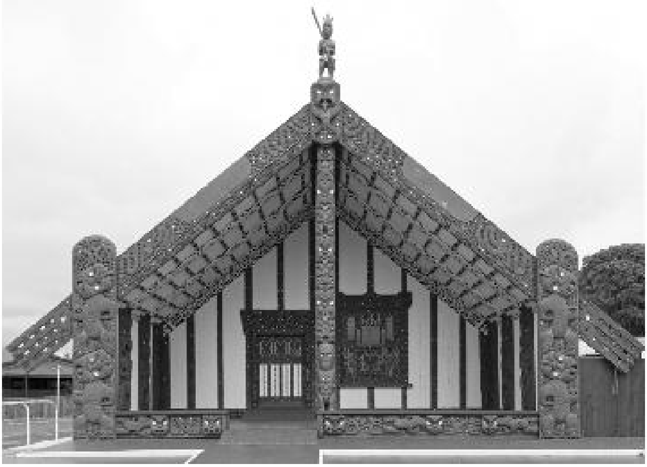


Figure 1.8

A marae

Māori is showing an increasing number of English words in its vocabulary (often, these refer to concepts and objects that have been invented in more recent times); however, the borrowing is a two-way street. A distinctive feature of New Zealand English has become the increasing use of Māori loanwords for names of places, people, organisations, words relating to sociocultural concepts and objects (such as marae and others just mentioned), and also the names of flora and fauna (though, interestingly, kiwi is not among them).

Russian

[In answer to the question ‘Which of the languages you speak do you consider the most beautiful?’] My head says English, my heart, Russian, my ear, French.

Vladimir Nabokov (1964), interview, Life magazine

Russian is a Slavic language within the Indo-European family; it is therefore a distant relative of English and German, and so has similarities in basic vocabulary such as kinship terms and numerals; Russian brát and trí; English brother and three; German Bruder and drei). Russian is spoken as a first language by roughly 160 million people, located mainly in the Russian Federation. There are also large Russian speech communities in Ukraine, Moldova, Latvia, Lithuania, Estonia and Israel, which makes it the most widely spoken of the Slavic languages. In this book, we have decided against presenting Russian examples in the Cyrillic script and have opted instead for a version of the Roman alphabet to enable those of you not familiar with Russian to gain a rough idea of the pronunciation.

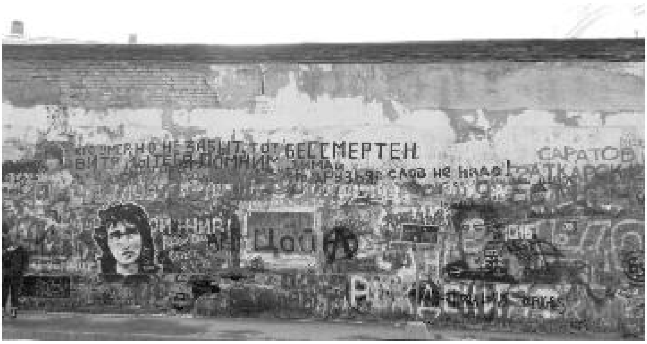


Figure 1.9

Tsoi wall in Arbat district of Moscow. Fans of the Soviet musician Victor Tsoi write on this wall in remembrance of their hero.

Russian is an inflectional language, showing many important grammatical distinctions, such as changes to the form of individual words. For example, Russian nouns and pronouns in the language distinguish three genders (masculine, feminine and neuter), and as many as six cases (these are called nominative, accusative, genitive, dative, instrumental and prepositional). Case marking such as this allows for much freer word order (the endings make it clear who is doing what and to whom).

Another rather famous feature of Russian is something called aspect (this is actually a translation of the Russian word víd). Aspect tells you things such as whether the event is well and truly over or whether it continues for a period of time. For example, Viktor čitál roman and Viktor pročitál roman both translate as ‘Viktor read the book’, but in the second example it means he read it to the end (that is, finished it). In English, up sometimes carries this meaning, but not consistently (compare I ate the cake and I ate up the cake).

The Russian lexicon is mostly made up of Slavic words that are shared quite widely through the family of Slavic languages. In addition, Russian has borrowed vocabulary from many European languages, particularly French, which was the language of the Russian court throughout the eighteenth and nineteenth centuries. Russian has also been a source of loanwords in languages such as English and German. Loans into English from Russian include beluga, mammoth, Soviet, tsar and vodka – in fact, most of our case-study languages have contributed in some way to the lexical coffers of English!

Turung

Let us all join forces to promote linguistic diversity and multilingualism as a key element in our efforts to build a better world and a life of dignity for all.

UN Secretary-General Ban Ki-moon, International Mother Language Day 2014

Turung (like Mali and Gurindji) is spoken by a fairly small speech community, with around 1200 speakers in six villages in Assam in north-east India. There are five different language families represented in this part of India. Turung is a Tibeto-Burman language, but it has a large set of Tai words in its lexicon.

Turung is not a language of wider communication, and in families in which the father speaks another language, the children typically do not learn Turung from their mothers. Other changes that reduce the strength of Turung include the fact that fewer and fewer domains of use remain for Turung; marriage with people from other language communities, the use of English and Assamese in education (there is no tradition of school-based learning in Turung), ever-increasing publications in Jinghpaw (from Burma/Myanmar) and the need for educated speakers to relocate outside the villages for work all pose major threats to the continued survival of Turung.

Figure 1.10

This photograph shows a group of Balipathar people who worked with Stephen Morey during his research on Turung in 2004. The principal consultant, Mr Aishu Shyam, passed away in 2011.

Like many of the languages of Southeast Asia, Turung is a tone language. Basically, tone is used along with specific vowels and consonants to produce distinctive sounds and words, and for Turung this means that there is a specific pitch associated with each syllable in the language. There is a three-way contrast for tone, marked by the superscript numbers ^{1,2} and ^{3}. Compare the following three words, which differ only in their distinctive pitch levels: waa1 ‘the grammatical category of definiteness’ (falling tone); waa^{2} ‘return (high rising then falling tone); waa^{3} ‘tooth’ (level).

The word classes in Turung include nouns, verbs, nominal and verbal particles, and clause markers (there is no separate class of adjectives; these are included within the noun class). A notable feature of its grammar is that it is an isolating language. This means that the words don’t change their shape but are invariable; in other words, there is basically no inflection in the language (nothing in the way of endings as in Japanese or Russian), and each word expresses a single semantic or grammatical value.

Our material on Turung is drawn from a grammar written by Stephen Morey. Our thanks to Stephen for his willingness to share the Turung grammar with us.

Further reading

Charles Hockett (1960) set out his design features of language and later expanded them with his co-author Stuart Altmann (1968). On the distinction between the signifier and the signified, see de Saussure (1966). On potentially infinite meanings, see Lyons (1977: 78). On the topic of animal communication, see von Frisch (1967); see Holldobler and Wilson (1977) on insects that work in colonies, and Konishi (2010), and Liu and Kroodsma (1999) on songbirds. Bolinger (1949) has a lovely discussion about the non-arbitrariness of the sign, and Benczes (2019) looks at people's creative language use and their search for meaning in sound. We draw on Johnston and Schembri (2007) in our discussion on the non-arbitrariness of signs in Auslan. On the examples of Irish English discussed in this chapter, see Filppula (2008). If you are interested in writing systems and the history of writing, we recommend Coulmas (2002), Rogers (2004) and Allan (2010). Aitchison (1992) presents a lovely discussion of Nim Chimpsky. Taylor (1990) describes how the debate between prescriptive and descriptive approaches to language has been played out for more than 300 years (beginning with the philosophers John Locke and John Horne Tooke). The distinction between competence and performance was established by Chomsky (1965). We recommend that you view the whole 'Concerning language' interview between Stephen Fry and Hugh Laurie (https://www.youtube.com/watch?v=ZFD01r6ersw).

A useful preliminary resource on many of the topics we cover is Crystal (1997). Chafe (1994) provides an exploration of the ways we use language to talk about distant times and places, and looks at how this impacts our thinking. Kenneally (2007) presents a recent review of research on the origins of language, with side notes on relationships among researchers across the field.

Basic data on the major languages from our case-study set (English, German, Japanese and Russian) were drawn from relevant entries in Brown (2006). The Ethnologue website provides numbers of speakers of different languages (https://www.ethnologue.com).

Crystal (2004) presents a very readable overview of the history of English, and books on English as a global language, and English grammar and its usage are in abundant supply. We also expand on Pennsylvania German using Burridge (2007). Additional material on the history of Japanese was drawn from Frellesvig (2010). An immensely readable account of Russian and Japanese structure and social life is provided by Comrie (1987) and Gottlieb (2005) respectively; these chapters also have useful additional references at the end. Language examples are drawn from a range of sources, including online dictionaries. We have drawn our material on Auslan primarily from Johnston and Schembri (2007). This is an accessible introduction to both Auslan and to sign language linguistics. Chapter 3 in that book includes a survey of sign language in relation to Auslan, including a brief history of New Zealand Sign Language (NZSL). More information about NZSL is available in Kennedy et al. (1997) and Kennedy (1999). Our main source for Gurindji is Meakins et al. (2013c), and for Gurindji Kriol, Meakins (2013a) (which contains references to other useful materials). A discussion of language practices in Gurindji-speaking communities is available in Meakins (2008). Most of the material on Mali is drawn from Stebbins (2011); additional materials relating to Mali are available in Stebbins (2004), Stebbins with the assistance of Tayul (2009, 2012), and Stebbins and Planigale (2010). For Māori, we have relied mainly on Bauer (2003) and Harlow (2007); additional research on Māori is published in various issues of the journal Te Reo (https://www.nzlingsoc.org/journal). The materials on Turung are drawn from Morey (2010). In addition to this grammar, Morey has published several papers of specific aspects of Turung and neighbouring languages, with a particular focus on tone, nominalisation and language contact. They are listed in the bibliography of Morey (2010).

Exercises and discussion points

1 Linguistics terminology

In this chapter, we have developed specific terminology relating to aspects of language; however, many of the terms used in linguistics also have everyday meanings that are used widely. Survey some friends, family members or acquaintances to find out what meanings they assign to the following words:

competence, gesture, grammar, language, paradigm, sign, symbol

How do the meanings used in linguistics relate to the more everyday understandings of these words?

2 Ideophones

Below are some ideophones from Japanese. The list contains: a number of forms that refer to different experiences of pain, two forms that describe elevated heart rates, and other expressions relating to body states and processes. Can you guess what the meanings are? What are you using to guide your thinking? To what extent do these words support or undermine the notion of the arbitrariness of the sign?

(a) paku paku

(b) bori bori

(c) doki doki

(d) gachi gachi

(e) hiri hiri

(f) chiku chiku

(g) zuki zuki

(h) shiku shiku

(i) sowa sowa

(j) zoku zoku

3 Language productivity

Using the pronoun paradigm in Table 1.3 and the sentence frame from example 1.4 (repeated below) see how many different sentences you can create.

(1.4) [A] locked [B] out of the house accidentally last week.

4 Style questions

Writing styles change over time. Read the following passage by Henry James from the preface to his novel The Portrait of a Lady, first published in 1881, in which he talks about procrastinating in Venice. Rewrite the passage in a more modern style. What changes did you make? What is gained or lost in the process?

I had rooms on Riva Schiavoni, at the top of a house near the passage leading off to San Zaccaria; the waterside life, the wondrous lagoon spread before me, and the ceaseless human chatter of Venice came in at my windows, to which I seem to myself to have been constantly driven, in the fruitless fidget of composition, as if to see whether, out in the blue channel, the ship of some right suggestion, of some better phrase, of the next happy twist of my subject, the next true touch for my canvas, mightn't come into sight.

5 What your teachers really say

Carefully record two minutes of your lecturer's speech in your next class. How does their speech compare with a written passage on the same topic? Were there any features of their speech that were difficult to write down? What was happening as they spoke that may have impacted their fluency?

Research project

In this chapter, we described the descriptive and democratic approach that now characterises modern dictionaries. By comparison, early dictionary entries often reflect the personalities of the dictionary makers – their likes and dislikes, tolerances and intolerances. Samuel Johnson's Dictionary of the English Language (1755) contains a number of entries that are now famous. One of the best known is the definition of the word oats: 'A grain, which in England is generally given to horses, but in Scotland supports the people'. Later, Johnson admitted to greatly enjoying this swipe at the Scots. He also had an array of derogatory labels for any word that he judged to be wanting in some way.

Johnson's dictionary is now available online. Investigate Johnson's entries of the following words and answer the questions below.

lexicographer, tory, newfangled, finesse, trait, ruse, mighty, sensible, hyper, to colour, dumbfound, clever, coax, chum, horrid, mishmash, flippant, fuss, banter, simpleton, shabby, chitchat, frisky, bamboozle, nowadays, spick and span, abominably

Note that in order to access the entries for these words, you will need to consult the scanned version of the dictionary using 'page view'.

1 Try to work out why Johnson branded certain words, and what the labels (such as 'low' and 'barbarous') signified.

2 Compare Johnson's accounts of these words with the entries you find in a modern collection (such as the Oxford English Dictionary or the Macquarie Dictionary).

3 How impartially descriptive is the modern dictionary compared with Johnson's work?

4 Now examine the entries for these words in the web-based Urban Dictionary and describe the approach of this dictionary. What do you think might be the value of a dictionary that is regulated by volunteer editors and rated by site visitors?

For more activities, please go to www.cambridge.edu.au/academic/linguistics.

2

What linguists do

◆

> Training in linguistics is training to think creatively, associatively, and in systems. We linguists see beneath the surface, we see solutions to puzzles, we see patterns in chaos, we are not afraid of ambiguity – all of which can (and should) be brought to bear on the challenges and needs of our world. We can and should find wide application of our powerful analytical skills, systems-thinking and problem-solving.
>
> Interview with Anna Marie Trester, 'The Career Linguist', 15 May 2017, *Superlingo*

2.0 Introduction

As we argued in the introduction to Chapter 1, it is through language that we transmit ideas, facts, opinions and emotions to each other. It is also through language that we conduct our social lives. So, as both the building blocks of communication and the mortar for sticking us all together, language deserves to be studied. Linguistic awareness should form part of general knowledge – alongside such domains as basic maths and science, current affairs and physical health.

But what about those highly specialised bits and pieces of linguistic knowledge we present throughout this book? Where in later life can you apply the facts you learn about how languages work – facts to do with, say, phonetics and phonology (sounds), syntax (grammar) or semantics (meaning)? Perhaps even more importantly, how can you take the skills and approaches developed through studying language out into the wider world to apply in your working life? As well as pattern finding, problem solving and systems thinking, this field of study includes the ability to observe social interactions and think about things from different perspectives. In ways you may never have considered, linguistic knowledge ends up having all sorts of applications in the workplace and is useful in a variety of jobs, both inside and outside the educational setting. In this chapter we introduce some of the ways in which you can press this knowledge into service. Each of our examples is illustrated with the 'voice of experience'– people describing how they have made the most of their linguistic knowledge. We hope you will be inspired to know that, in all likelihood, the things you learn from this book will prove useful to you in many other areas of your life.

2.1 Everyone should know about language

> Language, the most valuable single possession of the human race.
>
> Charles F. Hockett (1958), *A Course in Modern Linguistics*

Language belongs to everyone and we all need to pay more attention to the many ways in which it affects our lives, as well as the lives of others – particularly where it contributes to official smokescreens, deviousness in advertising and propaganda, linguistic discrimination, and manoeuvring in politics and business. As Dwight Bolinger once argued, it should be as natural to comment on the linguistic probity of courts, advertisers, businesses and offices as it is to comment on their financial probity. There needs to be more talk about language and it should be more informed talk.

As you will see in Chapter 15, the views often aired about language – particularly regarding language and social values – can be ferociously passionate and confident, but they are often woefully lacking in terms of the norms and standards of evidence we have come to expect of public debate on other topics. These days, most speakers of English probably know more about the ins and outs of climate change and stem cell research than they do about the ins and outs of their own language. As sociolinguist Ronald Wardhaugh puts it, 'Many educated people know more about space and time, uncertainty, and quantum effects than they do about nouns and verbs.' Yet, people feel they are experts simply because they speak a language. Many feel free to voice an opinion and will not budge from this opinion, even when it flies in the face of well-argued, rational linguistic evidence.

People can often be heard making negative judgements about others who use vocabulary, grammar and accents that they view as 'bad', castigating such people as 'uneducated' or even as 'stupid'. This behaviour seems out of place in an era where the basic human right of respect is understood to mean that people can no longer speak of or to others in terms that are considered insulting and demeaning. Yet, somehow this expectation does not extend to the way people talk about the language skills of others. Linguistic prejudices are usually accepted without challenge, and conscious and unconscious discrimination against speakers of non-standard dialects and low-status accents is rampant.

There are many areas in which you can find examples of the disastrous consequences of such linguistic bigotry, such as when popular (or ‘folk’) views on language are used to inform decisions that go on to affect the life chances of others. Sociolinguist Diana Eades, for example, has shown how ignorance about the different ways in which people speak can even end up putting the wrong person in prison; in Chapter 10, you will see how linguistic evidence has played a crucial role in dispensing justice in such cases. More recently, Eades has been exposing flaws in the methods used by immigration authorities to decide whether or not those seeking asylum are genuine refugees; the battle here is to ensure that well-researched principles of linguistics are put above the dangerously inaccurate folk linguistic views about how people speak (or should speak) that so often inform these methods.

2.2 Linguistics at work

Many careers call for clear and effective communication, and an understanding of the sorts of tensions and problems that can arise from human interaction. Many more careers call for the ability to navigate complex data in search of patterns, or to analyse and address problems within social or business systems. Those who are trained to examine how language works become more skilful in handling their own language and are better equipped to conduct their lives more effectively in the public sphere; they develop expertise in research, data management and creative problem-solving.

Below are just some of the wildly different areas in which linguistics graduates have found careers:

- Aboriginal affairs
- acting instruction or voice coaching
- administration
- advertising
- audiology
- captioning
- communications and signals

- computers, engineering and technology
- creative arts
- editing and publishing
- education
- field linguistics
- foreign affairs
- immigration and ethnic affairs
- information technology
- journalism
- language planning and policy work
- language teaching
- law
- lexicography (dictionary making)
- media
- social work and counselling
- speech pathology
- teaching English as a second or foreign language
- technical writing
- translating and interpreting.

We emphasise that this list is by no means complete and for more ideas we encourage you to visit some of the websites we provide in the online materials for this book.

Graduates in linguistics obtain jobs in these and many other types of careers. Some even go on to further study and specialisation in the various sub-fields of linguistics to become linguistic researchers and/or teachers themselves (otherwise they lead quite normal lives!). In the following sections, we will fully explore some of the most significant professions for which a knowledge of language is valuable and perhaps even essential.

Linguists in the legal system

Since the 1990s, the label 'forensic linguistics' has increasingly been used to describe the various applications of linguistics to real-world problems within the law. It is becoming increasingly common for linguists to be drawn into this area. They provide research on how the nature of a lawyer's questions can influence witness stories, or give research which shows the unequal treatment that speakers of non-standard dialects can receive. Some design and deliver training programs for investigative interviewing. Others provide expert testimony in the courtroom. Phoneticians might be called on for speaker profiling and voice comparison (comparing, say, the taped recording of a telephone conversation with the defendant's speech). Sociolinguists might be called on to determine whether the language attributed to an individual is an accurate representation of their speech (as in the case of fabricated or coerced confessions), or they may be asked to perform handwriting and stylistic analyses of documents to develop writer profiles to determine authorship (establishing, for example, whether it was the deceased who actually wrote the suicide note). Linguists are often called on to provide expert opinion when a trademark infringement has been alleged. Brand names often become household words (for example, *to hoover* now means to vacuum) and while some corporations don't mind this, many do, and most mind tremendously when other corporations move in on what they see as their trademark territory.

There are all sorts of applications of linguistics to the legal profession and plenty of other contexts in which linguistic knowledge can be applied. The Plain English movement is a good example. Language in legal settings is characterised by a highly technical vocabulary and difficult syntax, and this has been the case for some time. 'Whence does it come to pass that our tongue, so simple for all other uses, becomes obscure and unintelligible in wills and contracts?' What Michel de Montaigne wrote in one of his essays in 1580 still holds equally today. Lay people are often unable to understand the documents in which their legal and professional rights are set out. Expert evidence inside and outside the courtroom, police warnings, and even jury instructions have been shown to be notoriously difficult for non-lawyers to follow. Since the late 1970s, there have been social and political movements pushing for clear and simple language, and for some time now, linguists have been helping to translate legal language into a form that is more accessible to more people.

Misunderstandings in courtrooms can also result from linguistic and cultural differences of another kind. A number of linguists have been involved in the education of judges, magistrates and lawyers to expose the complexities of court interpreting, and to improve interpreting services in the legal system. Others work hard to expose the kinds of miscommunication that can arise when defendants speak a non-standard form of English (perhaps even a creole that is lexically related to English). A court might be dealing with a separate linguistic variety that is not properly recognised or understood outside a particular speech community. Misunderstandings in such contexts can have serious consequences and we provide examples of this later in the book.

Linguistics exposes miscarriages of justice

William Labov is an enormously important figure in linguistics, especially the field of sociolinguistics, and you will read more about him later. Here he describes one of the most fulfilling moments in his linguistic career.

> In 1987, I had another opportunity to test the usefulness of linguistics on a matter that was vital to a single person. A number of bomb threats were made in repeated telephone calls to the Pan American counter at the Los Angeles airport. Paul Prinzivalli, a cargo handler who was thought by Pan American to be a 'disgruntled employee,' was accused of the crime, and he was jailed. The evidence was that his voice sounded like the tape recordings of the bomb threat caller. The defense sent me the tapes because Prinzivalli was a New Yorker, and they thought I might be able to distinguish two different kinds of New York City accents. The moment I heard the recordings I was sure that he was innocent; the man who made the bomb threats plainly did not come from New York at all, but from the Boston area of Eastern New England. The problem was to prove this in court to a West Coast judge who could hear no difference between Boston and New York City speech!
>
> All of the work and all of the theory that I had developed since Martha's Vineyard flowed into the testimony that I gave in court to establish the fact that Paul Prinzivalli did not and could not have made those telephone calls. It was almost as if my entire career had been shaped to make the most effective testimony on this one case. The next day, the judge asked the prosecuting attorney if he really wanted to continue. He refused to hear further statements from the defense. He found the defendant not guilty on the basis of the linguistic evidence, which he found 'objective' and 'powerful'.
>
> Afterwards, Prinzivalli sent me a card saying that he had spent fifteen months in jail waiting for someone to separate fact from fiction. I have had many scientific results where the convergence of evidence was so strong that I felt that I had laid my hands on the reality behind the surface, but nothing could be more satisfactory for any scientific career than to separate fact from fiction in this case. By means of linguistic evidence, one man could be freed from the corporate enemies who had assailed him, and another could sleep

soundly on the conviction that he had made a just decision.

Linguists in education

> My linguistics training has been crucial. English Language draws on a lot of my linguistics study, and also keeps me in touch with how people think and write about language. It allows me to bring these ideas to young people, something really special considering when I left high school I had no notion of what linguistics was other than knowing languages and maybe writing grammar books on them. Not only this, but it has given me a very tolerant and descriptive approach to language generally, which helps me lead in-class discussions and educate future generations about the complexity and diversity of language (especially through the subjects I happen to teach!). In French it also has equipped me with some interesting methods of teaching an inflecting language that has a long tradition of being taught in certain ways.
>
> Interview with Jason, high school teacher (French and English language), 19 October 2015 on *Superlingo*

Traditionally, a good proportion of linguistic graduates enter the teaching profession, especially foreign language teaching. Being trained to think about a language in a scientific way is very different from the sort of unconscious knowledge a person has by simply being a speaker of a language, and this has all sorts of practical advantages in language teaching and learning. For example, knowing how English grammar and pronunciation work is particularly helpful for speakers of English who find themselves having to teach the inner workings of other languages, or having to learn them themselves. This is especially true when adults are involved (and adults are often more interested in foreign language learning than children, especially if another language is required for work). Acquiring a language later in life is far easier if learners already come equipped with some explicit understanding of how languages operate.

But studying other languages involves much more than learning words, their meanings, their pronunciation and the order in which they occur in sentences. Being a competent speaker of a language also means knowing how close to someone you can stand without causing them discomfort, how much to look that person in the eye, when to talk and when not to, what to talk about and how to talk about it, when to interrupt, how much to talk, when and how to compliment someone, and so on. Current research into **communicative competence** has shown that even speakers of the same language can have vastly different interactional styles. It is not surprising that discourse and conversational issues have come to form an important part of language teaching.

There are all sorts of ways in which linguistics can contribute to the education setting. At a very practical level, an awareness of different discourse styles can make for more effective teaching and student engagement in the classroom. And there are also many fairly specialised applications of linguistics to teaching – for example, in reading, composition and literature. The fields of sociolinguistics and discourse analysis have particularly important applications. Classrooms now contain children from vastly different linguistic and cultural backgrounds, and this can place them at a disadvantage if teachers are ignorant about concepts such as dialect, accent and standard language. Children might speak a language variety that is valued at home but marginalised in society, and if their ways of speaking are not valued as foundations for further learning, the outcomes for them can be catastrophic, with consequences well beyond poor school performance. Linguists have done much to draw attention to the complexity of non-standard language varieties, and have shown the educational benefits of a bi-dialectal policy in education. Teaching the school language will be much more effective if lessons build on what the child already knows about their home language. Approaching this home language as a legitimate and rich variety, and not as a collection of endless mistakes, will help a child to learn the standard language more effectively. Linguistics teaches that children can easily develop two accents, two vocabularies and even two dialects – one for school and one for home. They can also become skilled at switching between varieties according to context.

‘Thy true self’: Recognising non-standard varieties in the classroom

Below, linguist Jeff Siegel discusses his book *Second Dialect Acquisition* with Maria Zijlstra, presenter of the ABC Radio program *Lingua Franca*. She has just suggested that these days people are much more accepting of non-standard versions of different languages.

> Exactly. People used to think that anything except the standard was incorrect or broken or showed stupidity or anything like that, and although that still does occur … in Australia a lot of teachers think that Aboriginal kids just are too lazy to learn and speak standard English, and it’s quite disgraceful that the awareness about the variety of language that you say is changing among the general population hasn’t changed enough amongst the educational profession and you still get teachers making these amazing value judgements against their students. They talk about kids having

careless habits rather than speaking a different dialect, whereas if they spoke a different language, teachers would be a lot more tolerant.

But, as you say, I think the view is changing very slowly, at least with regard to national dialects, because people do hear all these different varieties with globalisation, they hear all kinds of different things.

Linguists and communication

> In user research, I do a lot of things I have done in linguistics research, from card sorts (where you get people to sort things into categories that make sense to them, used in semantics research) to ethnographies. I describe how systems are used to people who want to build software and digital solutions and make recommendations for that solution that will be easy for users to use as well as delightful to look at. If the business wants to do something with the software that won't resonate with users, I am the users' voice in advocating for the design that will be most effective for them while still meeting business needs.
>
> Interview with Abby Bajuniemi, user experience (UX) researcher, 26 June 2017, *Superlingo*

There are tangible everyday advantages to knowing how language works and how people communicate with one another. Linguistics has proven to be a great preparation for a range of roles where someone is needed to interpret the perspectives of one group into the frameworks used by another group – even when everyone seems to be speaking the same language. Where there are issues of cultural clash resulting from unstated assumptions, linguistics provides a framework for looking behind the surface of what people say, to consider the underlying messages and develop strategies for bringing this important information out into the light. This is useful in fostering communication across a wide range of settings, from business design to health communication and into public education programs of all kinds.

Linguistics and service provision for the Deaf

Louisa Willoughby is a sociolinguist who focuses on the intersecting areas of language and identity, and language policy/service provision for speakers of minority languages, particularly in health and education settings.

Since 2006, I have conducted research on various issues to do with service provision for Deaf people. These include projects on Deaf migrants, aged care service provision and analyses of Auslan users in the 2006 and 2011 censuses. I believe that it is imperative that my research have a connection to practice and have worked with the Deaf societies in each Australian state, as well as the Victorian Deaf Education Institute, Department of Human Services and the Deaf Connected at North Metro TAFE. These research projects have had a clear impact on policy and funding provisions (such as a funding extension to the Auslan fire alarm scheme, revamping of Auslan teaching in the TAFE sector and locations of the Video Relay Interpreting service). My research in these areas has formed the basis for the Victorian Deaf Society's submissions to three government inquiries and I appeared as an expert witness at the Senate Community Affairs Inquiry into Hearing Health in Australia (2009).

My research is fundamentally applied, but also engages in the major theoretical and academic challenges of the day. In my Australian Research Council project, for example, we are building innovative software that will give real-time feedback to Auslan learners on whether they produced a sign correctly; we are also comparing the efficacy of face-to-face and computer-assisted language teaching, and studying the ways in which students learn and retain sign languages (as distinct to spoken languages). My project on Deafblind communication similarly will answer a number of questions about how language is adapted to be conveyed through touch, while also producing better interpreter training resources. We are collaborating with the National Institute of Informatics in Japan to gather data on Japanese Tactile Sign Language as well as Tactile Auslan, so our findings will be relevant to the Deaf communities in both countries.

It is also often the case that when we are trained to think about how language works (for example, about sentence patterns and constructions), we become more skilled in managing our own language. This is particularly true when it comes to writing (which is not a natural activity in the same way that speech is). This is not to say that linguistic awareness will automatically turn us all into wonderfully effective communicators – there are enough examples of shoddy performances by linguists to know that this is not the case. An ignorance of linguistics won't prevent us from becoming effective communicators either. If that were the case, William Shakespeare might well have ended up a glove-maker like his father, John.

Nonetheless, research evidence supports the beneficial links between language awareness and language skills, especially when it comes to writing. After all, written language has to be explicitly taught after basic grammar, sounds and vocabulary have been acquired. With an awareness of language, and especially a feeling for sentence patterning, we can better evaluate the different choices that confront us when we draft something written, such as a speech or a report.

There are many groups of people who end up having careers in this area. There are, of course, authors who produce material for print and online media. There are technical writers, who specialise in producing materials such as instruction manuals and software documentation. And there are editors, who appraise and select content for publication. These are the people with keen eyes for ill-chosen words, grammatical and punctuation errors, and infelicities of style. Another related group you might not have thought about comprises those who work with dictionaries (lexicographers) and style guides. The development of good reference materials such as these requires qualified linguistic consultants.

Linguistics and effective communication

Linguist Richard Hudson has for some years worked hard to build bridges between academic linguistics and the wider community, and accordingly has written much about the applications of linguistics to real life. In a talk given to the workshop Identifying, Teaching and Assessing Key Skills in Linguistics, he outlines a number of life skills that grow out of an awareness of language. Here he expands on two that relate to communication.

> **Attention to form**. This is a specifically linguistic skill, but important in life to the extent that linguistic form matters – in short, very important. Maybe linguists have a pathological concern for form – for ‘interesting’ pronunciations, grammatical structures or whatever – but much of the population has an equally pathological disregard for it, at least at a conscious level. This disregard is pathological because most professional careers involve the production of written documents, and most written documents benefit from the attention to form that we associate with professional editors and proofreaders … Our students should be in a happy position somewhere between the two pathological extremes, able to pay attention to linguistic form when relevant, and not at other times.
>
> **General communication skills**. It would be comforting to think that linguistics makes a student better at communicating, and … this is probably the case. The trouble is that we know better than most that these two skills are only a small part of a complex package, where a strength in one corner may easily be offset by a weakness in another.

If studying linguistics *per se* made one a better communicator, then linguists should all be superb communicators, but none of us would seriously claim that this was so – we spend too much time reading opaque articles and listening to appalling conference presentations where the silly conventions of academic life offset the supposed insights of linguistics. Still, I think it might be fair to claim that linguists are quite good at communicating and that a good degree in linguistics is something of a guarantee of good communication skills.

Linguists in the field

There are all sorts of reasons for linguists going into the field to work in language communities. Sometimes it is a matter of working closely with the speakers to figure out, as much as possible, how a previously undescribed language works. The community will have certain expectations of the results of this collaboration; for example, the production of dictionaries, grammars, storybooks, educational DVDs and other pedagogical and community materials. In many parts of the globe, this is urgent work. Around the world, languages are dying at distressing rates – Native American, Celtic, Austronesian and Australian Aboriginal languages, to name just a few. Language endangerment, and what is known as language maintenance, represent the most urgent challenges confronting linguists, and many are now involved in some way in fieldwork, language documentation and language revitalisation. Even in those places where local languages appear healthy, the high status of bulldozer languages, such as English, Mandarin Chinese and Bahasa Indonesian, suggests that their survival is precarious. UNESCO clearly identifies language survival as a high priority. According to some estimates, this century will see more than half of the world’s nearly 7000 languages disappear.

Field linguists also find themselves working alongside other bodies, conducting language surveys, establishing literacy programs and language policy, translating documents, advising on linguistic and cultural heritage, and so on. In Australia, these activities could involve organisations such as: the Australian Institute of Aboriginal and Torres Strait Islander Studies (AIATSIS), which is the national institution for information and research on the cultures and languages of Aboriginal and Torres Strait Islander peoples; the Aboriginal Areas Protection Authority, which is responsible for overseeing the protection of Aboriginal sacred sites in the Northern Territory; and the Aurora Project, which runs projects in the area of Aboriginal education and Aboriginal affairs. In New Zealand, linguists have been working closely with community leaders to promote Māori cultural principles and the speaking of the language. As we discussed in Chapter 1, the success of educational initiatives such as the *Kōhanga Reo* (literally 'language nest') for young children and *Kura Kaupapa Māori* (Māori immersion in primary schools) has seen these programs being used as models in other places throughout the world.

Experiences of an outback linguist

Eugenie Collyer graduated in 2006 with an Honours degree in linguistics. Here she shares her experiences as a language research officer in the town of Katherine in the Northern Territory, outlining her role as linguist in a community language centre.

> The goal of a community language centre is to document, maintain and revi-talise Aboriginal languages in the region. Some language centres also run an interpreter service for Indigenous languages. I worked as a full-time linguist, with my role split between managing and training the interpreters and doing some community language documentation. There are various forms of documentation, such as working on dictionaries or grammars of the language, and creating resources with the community for use in school language programs such as videos, CDs, flashcards and books. Linguists are sometimes involved in class planning for the language programs as well.

Eugenie goes on to describe some of her projects below.

> One current project is a Plant and Animal (Ethnobiology) Project. These projects are conducted in collaboration with Aboriginal elders and a biologist. The aim is to document all the Indigenous knowledge of plants and animals for a particular language group … Many elders are concerned that their language will be forgotten, so they are keen to have it written down to be used for learning language and bush medicine by the next generation. I have been working specifically on a Plant and Animal Project for Mangarrayi language, from Elsey Station area (the station in *We of the Never Never*) … For example, sugarbag (bush honey) is very complex, with lots of names for the various parts of the nest, wax, honey, different types of bees, etc. Without a biologist, the linguist would not know which part was which. Likewise, without the linguist, the biologist would not know how to write down the language. And of course, without the elders, there would be no knowledge to document in the first place.

Linguists in advertising and marketing

There is probably no need to emphasise the importance of language in advertising. For centuries, advertisers have been using language to inflate, bamboozle or mislead. 'Promise, large promise is the soul of an Advertisement,' wrote Samuel Johnson in one of his essays in *The Idler* (no. 41, 20 January 1759). These days you can find e-versions of Johnson's essay besieged by web banner and frame advertisements (promoting everything from golf clubs, televisions and open colleges to 'fat-killing' vegetables, vanishing creams and 'further titles by Samuel Johnson'). The technological revolution has come a flourishing of new marketing opportunities – with an even richer 'magnificence of promises'.

Advertisers will want to use pleasing words with strong positive associations to highlight the qualities of a product (or to conceal its faults). A word like *crafted* sounds so much better than *made* or *manufactured. Crafted* gives the impression that something has been produced meticulously and probably by hand. Dressing up the goods with verbal glitz is an important part of the 'large promise' of modern advertising. Of course, the more advertising there is, the larger the promise has to be. Companies specialising in advertising spend millions of dollars annually researching what kind of wording helps to implant desire in potential consumers. And this is where linguists come in.

Companies work with linguists when they are preparing sales campaigns or searching for slogans and appealing names for brand new products – names that are meaningful, catchy, distinctive, memorable, easy to say and legally protectable. Creating a name from scratch has the advantage of avoiding trademark disputes, but it also allows the company to trade on the perceived relationships between sounds and meaning. Linguistic research into sound symbolism, for example, tells us that voiceless stops (such as [p] and [k]) have connotations of compactness and speediness. This is useful to know if you are a company seeking promising brand names for a new car. Just the right assemblage, of consonants and vowels can create a certain impression or atmosphere, and send out a subtle signal to buyers looking for the product that suits their image. As Bob Cohen from the firm Lexicon asks, '*Clorox* versus *Chanel* – which is going to be the hard-working laundry detergent and which the new fragrance?'

This sort of information becomes increasingly important for products being sold globally. There are some famous cases of names that have flopped spectacularly in non-English speaking countries (for example, in the Spanish-speaking market, the car name *Nova* means 'doesn't go'!) If an Australian wine company wanted to promote its wine in China, what are the culturally appropriate metaphors that could be used in the tasting notes? ‘An alluring nose with coffee and chocolate combined with forest floor aromas’ might not carry the desired message for potential consumers outside the West.

What’s in a name?

In their 2011 article, business gurus Dan and Chip Heath delve into how linguists contribute to branding and marketing processes. As part of this, they outline the story behind the naming of the disposable mini-toothbrush Colgate Wisp, the brainchild of the company Lexicon. Most won’t recognise the company name but they will certainly recognise many of its brand names, such as BlackBerry, Dasani, Febreze, OnStar, Pentium, Scion and Swiffer.

Such billion dollar names aren’t produced on a whim. As the Heaths point out, ‘great names do not come from lightning-bolt moments’. The process is long, tedious and expensive. Lexicon CEO David Placek employed 70 linguists across 50 different countries for this job. These linguists worked independently to brainstorm the vowels, consonants and metaphors that would help create the perfect brand name – one that would boost sales, wasn’t offensive in any language, was actually available (so not already taken or too close to an already existing name) and, most importantly, would evoke just the right image. The toothbrush is tiny (small enough to fit into the back pocket of a pair of jeans) and in the centre of its minuscule brush head is a liquid-filled minty bead. No rinsing is necessary. Using the brush doesn’t release a big mass of foaming lather; it’s more a light rush minty freshness (to paraphrase the current advertising blurb). So, the linguists were after something that didn’t so much convey the idea of smallness, but rather lightness, gentleness and softness. In the end, out of their long list of words and phrases, the team hit upon Wisp. According to the Heaths, ‘It was the perfect association for the new brushing experience and it tested well; it’s not something heavy and foamy, it’s barely there. It’s a wisp. Thus was born the Colgate Wisp.’

Linguists in health sciences

Language, communication and literacy difficulties and differences are far

> more prevalent than most people would imagine, and their impacts have profound effects on people's lives. Unfortunately, lack of knowledge in the world about both areas means these impacts tend to be negative. I highly recommend reading up on language, communication and literacy difficulties and differences, and the social model of disability.
>
> Interview with Annie, speech pathologist, 16 July 2016 on *Superlingo*

Speech therapy, or the more general study of communication disorders, is concerned with people whose language faculty is somehow different from the expected. This can be due to developmental problems or to an illness or accident that damages a language faculty that had previously been fully developed. An important task for speech therapists is to develop techniques that can help these people to improve their communicative abilities. In order to study and describe language that is 'not normal', it is important to know what the structure of 'normal' language is, and to master the terminology used to describe it. The methods and scholarship of linguistics provide this foundation.

Linguistics offers a wealth of knowledge to assist in the better understanding of mental health. In recent years, trained linguists have been working alongside speech therapists, psychologists and other health professionals to improve awareness of the relationship between communication disorders and psychiatric disorders, concentrating on the areas of prevention and education.

Linguistics and mental health

Deb Perrott is a consultant psychologist and linguist. Here she explores the various ways in which linguistics can assist with psychiatric measures, especially by spelling out more clearly the link between communication disorders and mental health.

> An Australian community sample of depressed adolescents showed that they used poor pragmatic skills in daily communication compared with their non-depressed peers; these problems included difficulty with initiating, continuing and terminating conversations. They also reported difficulty with managing topic change and extending topics and described the use of stereotypic language as being easier than extending a topic of conversation, as it gave the impression that the speaker was being understood and it required less energy than adding to a topic. Maintaining eye-contact is also a difficult task for depressed youth. Clearly, such research leaves us to contemplate some of the issues that may arise for such youth regarding future employment and training opportunities.

Linguists like Deb are key players in the quest to support communication research and assessment tools that enable a greater understanding of mental health. The important first step is to acknowledge the link between communication disorders and mental health disorders, and to build on this research. From such research, we can further refine our application to assessments of communication that may highlight mental health issues.

There are many other ways in which linguistics can be used to improve the medical care of patients. One area of growing importance is doctor–patient (or practitioner–client) communication, especially addressing breakdowns in communication. Those with a knowledge of discourse analysis are better placed to identify the discourse tensions that can occur between health professional and patient, especially where different linguistic and cultural backgrounds exist. Misunderstandings can involve different languages or language varieties in the medical setting. They can also arise because of the different discourse styles of doctors and patients. The medical profession is full of scientific expressions that distance patients from healthcare workers, thereby working against the healing process.

There is plenty of evidence that words can be curative (the placebo effect); conversely, they can harm a person's health (the nocebo effect). The mind–body effect is something that many in the medical profession are now taking seriously, especially when it comes to breaking bad news, and linguists have shown that different ways of delivering information to patients can result in radically different outcomes; healthcare professionals need to be aware of the impact of the language they use when recommending treatment. Many linguists are now applying the results of their research on a very practical level, by providing support to medical professionals and associated service industries (such as aged care, palliative care and other medical facilities) in the form of communications skills training and information booklets, videos, and podcasts on how to discuss medical issues with those affected.

Linguists in business and professional services

> We collect thousands of opportunities from different sources across the web and make them much easier to search for … My work is largely about identifying patterns in these contracts … My work is moving into a direction where I'll need to learn how to use regular expressions and I'm very excited to see how that goes. My university didn't offer any

> computational linguistics classes, at least not if you weren't an upper-level computer science major. But the way that I think about language is very different from how my coworkers do. It's fascinating to me the things that others don't know about how to analyse language, I guess just because that's my content area. They know infinitely more than I do about business development and computer engineering.
>
> Interview with Caroline Davis, text analyst, 31 July 2017 on *Superlingo*

The business community and professional services firms are now abandoning their narrow selection criteria and are beginning to opt for more diversity in their graduate recruitment. They are increasingly recognising the benefits of a humanities education in the business world. Annette Gordon-Read, reporting in *Time Ideas* (19 June 2013), describes how many surveys now show that 'employers have expressed a preference for students who have received a broadly based education that has taught them to write well, think critically, research creatively and communicate easily'.

Linguistics is in an excellent position here. The discipline emphasises many different transferable skills, such as problem solving, the collection of data, the analysis and reporting of findings, and recommending approaches to the solution of problems. Linguistic graduates have found themselves well trained for careers in business – in financial institutions, in marketing, in personnel work and in industrial relations – and, of course, any career that calls for clear and effective communication, and an understanding of human interaction.

Business prizes a humanities education

In an article published in the *Sydney Morning Herald* in 2013, Peter Cai reports on the increasing trend worldwide for business employers to seek graduates from humanities disciplines, such as linguistics, rather than economics or business studies. He argues that modern business must be able to engage effectively with the diverse range of people who are involved in their businesses, both directly and indirectly. For example, in the mining sector, working effectively with local communities is as important as finding mineral deposits.

> Business needs people who can understand the world beyond balance sheets and mathematical models. There are too many people at home and abroad who are showing that diversity is good for business for us to ignore them.

He quotes corporate strategist Peter Acton (a former partner at Boston Consulting Group, an international strategy firm).

> The ability to help a client to think through a complicated problem, to be able to articulate the logic and to come up with some creative ways of looking at problems, I think all these things come from somewhere other than straightforward vocational business education.

Another application of linguistics that you may not have thought about is safety in the workplace. Language is not a precise notation like logic, but is plagued with vagueness, indeterminacy, variability and ambiguity. While it would be neither good nor possible to eradicate these features (you'll see why throughout this book), ignorance of them can have dire consequences. Think of the airline industry. A number of aircraft tragedies have occurred because of simple miscommunication. On 27 March 1977, two Boeing 747s collided on the airstrip. In this instance, the KLM pilot had radioed, 'We are now at take-off', with the understanding that the plane was taking off (a Dutch construction for 'flying' as 'at fly'). The tower controller had understood that the plane was waiting on the runway (he understood it to mean the plane was 'in the takeoff position'). For all sorts of reasons, ambiguities can arise in language and the results can be calamitous – for example, when the same word has different meanings (in avia-tion jargon, *hold* means 'stop what you're doing'; in ordinary language it can mean 'continue what you're doing') or when different words sound the same (such as *two* and *to*; for example, *cleared two-seven* versus *cleared to seven*). But while faulty communication is a linguistic fact of life, there is much that can be done, particularly in the development of technology, to help prevent such communication tragedies. American linguist Steven Cushing is someone who has examined dozens of clear-cut cases in which confusing language has led to aircraft disasters, and in his book *Fatal Words* he proposes practical solutions to avoid such incidents in the future.

An impressive example of how applied linguistic research can contribute to our understanding of workplace discourse and culture is the Wellington Language in the Workplace Project. Since 1996, a team of New Zealand linguists at Victoria University of Wellington has been researching different aspects of interaction (both business and social) between people at their places of employment. This work has accumulated a vast database of 1500 interactions in a diverse range of private and public organisations, factories and small businesses. Areas of investigation include (among many others) the nature of instructions and directives, the importance of humour (for instance, teasing) and small talk, the role of swearing, the structure of formal meetings, the function of email, male and female speech styles, sexism in workplace documents, miscommunication, and cultural differences in styles of interaction. As the group's website explains, a broad aim of this research has always been to improve workplace communication and help to create a better working culture. The team has published a wide range of material with very practical applications; for example, it has covered effective communication in presentations, workshops and seminars, and provided training for those (such as immigrant workers) joining New Zealand workplaces for the first time.

Linguists in the creative arts

Language is central to a range of creative enterprises, particularly those associated with writing and performance. Among the many people with linguistics training contributing to the creative arts are authors, playwrights, poets and translators. Some linguists have made careers in the film industry, where they use their linguistic knowledge to help train actors to produce certain accents or grammatical patterns in order to sound like an authentic native speaker of a particular dialect or other variety of language, or even to produce mistakes to sound like a real non-native speaker. And unlike in the past, when there was scant regard for linguistic realism (especially when it involved Indigenous languages), linguists are often called in to provide language advice to filmmakers on matters like multilingual communication and on aspects of the languages themselves.

One area for which linguists have become famous is the creation of artificial languages in science fiction films. It started with J.R.R. Tolkien, who constructed a number of languages for the fantasy worlds in his many books (the Elvish language family is the best known). Tolkien was an Oxford professor of philology (a precursor to linguistics) and he spent a lot of his time working on constructed languages, creating more than 20 (he also nearly became a codebreaker during World War II).

Tolkien's work – 'fundamentally linguistic'

Tolkien's work has become widely known through the movie adaptations of his writings. On 30 June 1955, he wrote a letter to his American publishers, Houghton Mifflin Co., in which he describes his work as 'fundamentally linguistic in inspiration':

> The invention of languages is the foundation. The ‘stories’ were made rather to provide a world for the languages than the reverse. To me a name comes first and the story follows. I should have preferred to write in ‘Elvish’. But, of course, such a work as *The Lord of the Rings* has been edited and only as much ‘language’ has been left in as I thought would be stomached by readers. (I now find that many would have liked more.) But there is a great deal of linguistic matter (other than actually ‘Elvish’ names and words) included or mythologically expressed in the book.

Constructed languages are becoming more usual in films. A number of you will be familiar with the Klingon language in *Star Trek* movies. Its creator, Marc Okrand, is a linguist with a penchant for grammar. The Na’vi language spoken by the inhabitants of the fictional moon Pandora in the 2009 film *Avatar* was also created by a linguist, Paul Frommer. Since the 2009 release of the film, Frommer has expanded Na’vi’s vocabulary to more than 1500 words and has also published the grammar; Na’vi, like Klingon, is now a learnable language. It was Francis Nolan, a lecturer in phonetics, who constructed the fragments of Parseltongue language for the *Harry Potter* films. The linguist and anthropologist Christine Schreyer served as a consultant for the Hollywood blockbuster *Man of Steel*, helping to develop a new language for Superman’s home planet of Krypton.

In fact, there are hundreds of people out there inventing languages. The *conlangers* – or ‘constructed language creators’ as they are known – have their own Language Creation Society, which, as its website describes, is for ‘conlangers, would-be conlangers, those interested in or curious about conlangs, and anything else to do with conlanging’. Not surprisingly, it was this society that HBO contacted when it wanted to flesh out both the Dothraki language and the Valyrian languages for *Game of Thrones*, the television adaption of George R.R. Martin’s fantasy novels; the society’s president, linguist David J. Peterson, got the job.

Linguists in computing and technology

> I don’t have a computer science background, and I’ve never been a big computer nerd. I’ve learned everything on the job, and it turns out I quite like computers after all. If you have the type of pattern-finding brain that linguists tend to have, chances are you might enjoy the technical side of things when it’s applied to something interesting. Developing some technical skills as a student can help you in many career paths, but it’s not an absolute necessity. If you’re looking for a place to start: regular expressions, bash, and Python are all very useful and relatively easy to

learn.

Sasha Wilmoth, Linguistics project manager, 30 October 2017, *Superlingo*

With speech-based applications now used increasingly in human–computer interfaces, natural language processing and speech recognition technology have become two very significant areas of information technology. Many trained linguists are now seeking jobs in computational linguistics, which involve work in such diverse areas as: speech technology, machine translation, information extraction, text mining, voice recognition and voice translation software, recording and archiving technologies, and software development for (field) linguists. The practical applications of this work are many: better speech-recognition platforms for telephony- based systems and computer avatars, computer aids for children with learning difficulties, enhanced perception for hearing aid and cochlear implant sound processing, as well as a range of security and forensic improvements (for example, speaker authentication and identification). Linguistic research is the key to progress in all these areas and training in linguistics is a necessity for these new career paths.

The Big Australian Speech Corpus

Professor Denis Burnham (2011) from the University of Western Sydney is talking here with ABC journalist Paula Kruger about the Big Australian Speech Corpus (the Big ASC project) (also known as AusTalk) – a huge research project that is currently collecting the voices of more than a thousand speakers of Australian English nationwide. Paula asks why a database is needed.

> Well it’s actually a really critical time in Australian history linguistically and otherwise. There’s a large amount of digital media coming from overseas … There’s been no large auditory and visual database collection so we need that in order to teach machines to understand and to speak Australian accents. So it would be useful in applications like relatively minor ones like ordering taxis and ordering pizzas over the phone and things like that so that they understand Australian accents but also more importantly things like cochlear implants and hearing aids. Computer programs for hearing-impaired children and for children with learning disabilities – we’re using the Australian accent to do those things as well.

Over the last few years, linguists and language technology researchers from around Australia have also been working together to create another massive online database of English in Australia; that is, a national collection of computerised spoken and written language, in all its forms and diversity, past and present. The Australian National Corpus, or AusNC, was officially launched on 26 March 2012, and is an ongoing project to create a living multimodal collection (audio and video files, transcriptions, written texts and so on) available to anyone with an interest in language and language use. We encourage you to visit the Corpus website and make use of this wonderful resource for some of your research projects.

Work in computing and technology has been important in the development of facilities for digital archiving of endangered language materials. A notable example is the Pacific and Regional Archive for Digital Sources in Endangered Cultures (PARADISEC), which has established a framework for accessioning, cataloguing, and digitising audio, text and visual material, and preserving digital copies.

2.3 Some final remarks

Linguistics is in an interesting position because it is basic to so many different disciplines, not simply those that have language as their main object of study. The subject therefore crosses over many of the traditional boundaries of arts, social sciences, science, technology and business. Because of this, you will find that you can combine linguistics in your studies with a whole range of quite different subjects to create an exciting degree program.

We've only been able to give you a brief sample in this chapter of the many different fields that require people with linguistic knowledge, but throughout this book we will be providing further tastes of the sorts of enterprising things you can do with the material you acquire in your introductory linguistics classes.

Further reading

For a collection of articles showcasing linguists at work see Oaks (2001) and the fantastic Superlingo blog by Lauren Gawne, which includes a collection of interviews with people doing amazing things with their linguistics training. We include a link to this in the online resources.

On a lack of knowledge about language, see Wardhaugh (2010).

For more on the dangers of poorly informed debate around correct language, see Burridge (2010). Our discussion of the various applications of linguistics was informed by several works, as referenced here: language and the law – Cohen (1995), Eades (2012), Kelly, Leben and Cohen (2003), Koch (1985), Langford (2000); language and safety – Cushing (1994); language and health sciences – Cordella (2004), Cordella and Musgrave (2009), Perrott (2014); language, business and advertising – Cai (2013), Heath and Heath (2011); language and communication – Householder (1971), Hudson (2001), Willoughby (2011, 2012, 2014); language and computing – Mistica et al. (2008); linguists in the creative arts – Carpenter (1981), Peterson (2015); different aspects of language and education – Hudson (2004, 2007), Labov (1969, 1972a, 1972b, 1972c, 1997), Siegel (2012), Trudgill (1975); the importance of language in people's everyday lives – Bolinger (1980). We also recommend that you visit the manifesto for LASER (Language Analysis in Schools: Education and Research), which highlights the central place that language analysis has in education generally (including subjects like science, geography, maths and history): see http://clie.org.uk/laser.

ABC Radio's *Lingua Franca* is now an archived program and recordings and transcripts can be found at http://www.abc.net.au/radionational/programs/archived/linguafranca.

Exercises and discussion points

1 'Proper' English and education

(a) If a child said, 'It wasn't me, David done it!', what would you think about their language?

(b) If this was said by a primary school child, what do you think their teacher's response should be?

(c) If a student wrote something like 'It wasn't Socrates, Plato done it' in an essay, would you expect the same response from the teacher?

(d) What does all this say about the role of education in language acquisition? What does it say about differences (and similarities) between speech and writing?

2 Car names – sound symbolic associations

Car companies seem to have a deliberate C-naming strategy. Make a list of as many car names you can think of, and divide these into existing words (and these will come with their own associations) and nonsense words (which may have sounds that convey meanings). Bearing in mind that the letters like <c> can symbolise various sounds, try to determine what the symbolic associations of these sounds might be.

3 Good and bad language

(a) Comment on your own opinion of the acceptability of the following pronunciations. Ask friends or family members what they think.

ask → aks	nothing → nothingk
known → knowen	vulnerable → vunerable
aitch → haitch	secretary → secatry

(b) Comment on your own opinion of the acceptability of the following lexical items. Also ask friends or family members what they think.

cookies, youse, movies, gotten, trash, ask (as in *a big ask*)

(c) Now make a list of examples of language usage (for instance, words, word meanings, pronunciations, grammar) you have observed others adopting that you think are wrong (or that 'get up your nose'). Include examples you have used that others might have criticised. Do these differences reveal anything about changes that might be underway in the language?

(d) What do you think is really going on when people object to words and word usage in this way? How do they justify their objections? Are they valid arguments?

4 Advertising in the global village

Branding can be risky business in a global market. Consider the following extract from Melbourne's *The Big Issue* (no. 65: 8).

> The influence of other cultures enriches and strengthens our own immeasurably, says David Suzuki. But try telling that to Clairol, who introduced a curling iron named the 'Mist Stick' into Germany only to discover too late that 'mist' is German slang for manure. Scandinavian vacuum manufacturer Electrolux used the slogan: 'Nothing sucks like an

> Electrolux'. Pepsi's 'Come alive with the Pepsi Generation' was translated into Chinese as 'Pepsi brings your ancestors back from the grave'.

Other types of problems are revealed by a decision a few years ago to change the name of the soft drink Pepsi to Pecsi in Argentina. Last year, advertisers in Spain followed suit and changed the name of the soft drink to Pesi. Imagine you are working for a brand-development firm. Identify some English brand names that may be problematic for speakers of some languages and describe why these names are problematic (think of issues to do with sounds, meaning or grammar).

5 Words of desire

Consider the words listed below. Each of them has positive connotations associated with certain types of food and negative connotations in another context. Identify what kind of food would benefit from being characterised in each way and then list another context (or food) in which the word would carry a negative connotation.

(a) chunky

(b) delicate

(c) tender

(d) addictive

(e) raw

Research project

Grammar awareness and writing skills

In this chapter, we have made the claim that there are beneficial links between language awareness and language skills, especially writing. But this is a controversial area. A number of early research projects that matched the writing skills of children, both with and without formal grammar instruction, found no positive correlation between explicit knowledge of grammar and writing skills. Research this claim and then write an essay arguing for one position or the other. A good place to start is Dick Hudson's website: http://www.dickhudson.com.

For more activities, please go to www.cambridge.edu.au/academic/linguistics.

3

Ways to study language

◆

> Language, eight-armed, problematic, demiurgic, infinitely entrailed, must be honoured. Its riddling, jokey, mischievous, metaphoric, flawed, lapsible, parapraxic life must not be repressed, but tolerated, pleasured, submitted to, enjoyed, and so revealed for what it is.
>
> Galen Strawson, *Quarto*, March 1982

3.0 Introduction

The challenge involved in studying language is that it is all around us all the time, so it can be difficult to gain some distance from the subject. Many of us spend large portions of each day talking with as much conscious attention as we give to breathing. As a result, reflecting on language can be a bit like asking a fish to think about water. How do you get outside it far enough that you can see what is there?

The concepts we introduced in Chapter 1 are part of a solution to this issue. They provide us with some basic tools for thinking and talking about language. This chapter sets out some of the research strategies that we can use as part of the process of answering questions about language. Of course, the types of tools and strategies you use will depend on the question you are trying to answer. (It's pretty hard to make an omelette with a hammer!) So, in this chapter we'll describe some approaches to research on language and the types of research questions that they can be used to explore.

It can be helpful to note that many techniques and research strategies widely used in other disciplines are also relevant to various specific types of research in linguistics. For example, **psycholinguists** study language in relation to psychological processes; not surprisingly, they employ a range of experimental strategies also used in psychology more generally. **Sociolinguists** study the relationship between language and society; they share many of their methods – including participant observation (where the linguist participates in the life of the speech community), observation, case studies, interviews, questionnaires and focus groups – with the disciplines of anthropology and sociology, and other disciplines with a similar focus on social values and patterns of interaction.

There are also some strategies for studying language that have been developed within various branches of linguistics that are not widely used in other fields. Within the general area of experimental research, the study of speech sounds and the workings of the speech organs (that is, the lips, mouth and especially the tongue and larynx) are part of a specialisation that touches on anatomy and physics, while specifically focusing on the human voice. Some of this research involves highly specialised equipment that is not used elsewhere. The production of dictionaries and grammars requires close study of the distribution and meanings of different language forms. This type of analysis requires orderly examination of the data and careful use of logical reasoning, using methods that are quite distinct from those in other disciplines, and and very good training in clear thinking!

In the following sections, we distinguish between different sources of linguistic data and the types of uses to which this data can be put. We close with a consideration of some of the ethical issues associated with linguistics research.

3.1 Sources of linguistic data

> Language never deceives, if only we know how to question it aright.
>
> Richard Chenevix Trench (1851), *On the Study of Words*

We can tackle the problem of stepping out of language far enough to see what it is doing in a number of ways. One is by developing collections of texts to look back over for patterns and clues about how language works. Another is by observing the language behaviour of other people around us. We can also ask people directly about their views on language. Each of these approaches has its own set of advantages and challenges, as described below.

Corpora

A **corpus** is a collection of language data that we can examine to find out about how language is being used or how it has been used in the past. (Note that it has a rather fancy plural, **corpora**, although the more mundane **corpuses** is also used.) In Chapter 2, we gave you a taste of two online corpora in Australia: the Australian National Corpus (or AusNC) and the Big ASC project (or AusTalk). Such corpora are collections of texts, by which we mean things people have written and also things that they have said that have been recorded and transcribed. Any collection of texts could be called a corpus, but mostly today within linguistics a corpus is understood to be a digitised collection of texts that is searchable and has been annotated to make it more useful.

People have always used texts as sources of language for analysis. There is a long tradition of studying the history of English, for example, and it relies heavily on the texts that have survived from the earlier stages of English. A wonderful online resource for the history of English is the *Helsinki Corpus of English Texts: Diachronic and Dialectal*; it has texts covering the period from 750–1700 ce and a corpus of dialects partly based on transcripts of interviews with speakers of British rural dialects in the 1970s. Within AusNC, there is also a historical collection, the Corpus of Oz Early English (the COOEE corpus). The total corpus is substantial: it contains just over two million words of Australian, New Zealand and Norfolk Island origin from the years 1788 to 1900. It has a combination of speech-based texts (speeches, plays, court proceedings, testimonies) and written texts (government legal documents, public documents such as newspapers, as well as private documents such as letters and diaries). The collection offers wonderful possibilities for the exploration of past forms and structures, and has opened windows on variation in the early history of postcolonial varieties of English.

I should not care a b***** b*****

A recent arrival on the corpus scene is the *Old Bailey Corpus*. This is a massive historical corpus of spoken English, comprising a selection of the almost 200 000 criminal trials held at the Old Bailey (London's Central Criminal Court) over a nearly 200-year period (1720–1913). It makes for fascinating reading.

The proceedings were taken down in shorthand in the courtroom by scribes, using quill, pen or pencil, and probably the pen-writing method of shorthand (until stenotype machines became available in the 1870s). Not only do these records give a picture of what went on in the criminal courts from the seventeenth through to the early twentieth century, they also provide rare access to the language of non-elite people and, most importantly, the spoken styles of those times. Even swear words (notoriously absent from other writing) make an appearance. Dating these expressions is always difficult on account of taboo; they can have a life on the streets long before they happen to make an appearance in print. The Old Bailey transcripts reveal, for example, that as early as 1821 the word *bugger* (though often in heavy linguistic disguise!) could be used to strengthen negative constructions (*It cannot be helped now, I should not care a b – dy b – gg – r if I was going to be hung up for* it). This is more than a century earlier than previously thought.

Text collections are an important starting point when a grammar is being prepared for a previously undescribed contemporary language. However, corpora can be used to answer sociolinguistic as well as descriptive questions. For example, Englishes all around the world have seen the emergence of innovative forms, such as the ‘quotative *like*’ (*And she was like, ‘Give me a break will you!’*). Corpus evidence can be used to identify the gender and age of a speaker, and check who uses this construction and how often it occurs for different groups. Linguist Celeste Rodriguez Louro is currently investigating the social and linguistic factors that are implicated in the use of this marker in Australian English, and she is using two corpora for her research: a collection of narratives of personal experience from the State Library of Western Australia oral history collection (1963), and the University of Western Australia (UWA) Corpus of English in Australia (2011).

Our brief description of some corpora shows that there can be a lot of variation with respect to the number of texts and overall size of the collection; for large corpora, this is often measured in millions of words, while in smaller corpora it may be measured in hours of recordings. Corpora also vary in their representativeness. Some cover a wide range of text types, and are broadly representative of the wider community in terms of age, gender, dialect and socio-economic group membership, while others are quite restricted in size, focusing on one particular group. The size and composition of a corpus typically reflects the resources that were available in compiling the corpus and the purposes it was developed to meet. In situations where languages are disappearing, it can be a matter of linguists salvaging what they can. So, a corpus of 50 hours of recorded narrative by just a handful of speakers – all men and all of whom are quite elderly – may be impressively large and significant if it involved trekking on foot to a remote location and making a record of a dying language. Clearly, such a corpus is fundamentally different from the corpora that have been compiled in dictionary projects for major languages using data-mining techniques and containing millions of words. Both are useful for different purposes.

Where the bloody hell are you?

In February 2006, Tourism Australia launched an international tourism campaign with a television advertisement showing images of everyday Australians set against a backdrop of famous landmarks, and concluding with the very Australian invitation, ‘So where the bloody hell are you?’ The advertisement managed to get itself banned from British television (it was also censored in North America). So, what light might corpus evidence throw on the international squeamishness surrounding the swear word *bloody*, described as the ‘great Australian adjective’?

Something corpora are very good for is counting things. The ART corpus (covering Australian talkback radio between 2004 and 2006) contains thirteen occurrences of *bloody* in 20 375 sentences (six per 10 000). In the conversational data of the ICE-AUS (the Australian contribution to the International Corpus of English), *bloody* occurs 46 times in 22 933 sentences (20 per 10 000), and in the ICE-NZ corpus (the New Zealand component) it occurs 43 times in 60 175 sentences (seven per 10 000). Finally, in the Corpus of London Teenage Speech, there are 291 instances of *bloody* in 107 429 sentences (27 per 10 000).

Although these collections of texts aren’t exact matches, the results are quite telling. We in the Antipodes lag well behind teenage Londoners in our use of *bloody*. Clearly something else was going on with the banning of this advertisement in the United Kingdom!

The process of recording, transcribing and marking up texts has important consequences for the information that can ultimately be retrieved from them. American linguistic anthropologist Elinor Ochs points out that an important feature of transcriptions is that they are selective: they do not represent everything about an interaction that could have been recorded. This process of making a selective representation of the interaction allows us to work with the data more readily, but it may come at a cost. Because transcriptions focus on those aspects of the interaction that are of interest to the researcher, they inevitably omit other information – some of which may have been relevant to the research. Ochs argues that it is important to make the selection of data a conscious and deliberate process; that is, one that is disclosed to the reader when research is later reported. The way language is represented is also important. Ochs refers to the loss of information about sound play in transcriptions of child language as a case in point. When researchers transcribe texts using Standard English spellings of words, it can later be difficult to identify examples of sound play.

In recent years, technological advances have helped to overcome these difficulties, with the development of computer-mediated communication, and improved possibilities for creating and disseminating high-quality recordings of language. All this has made for vast improvements in corpus design. Many modern corpora now aim to be multimodal: they contain not only text-based sources, but also high-quality recordings, both audio and video. New additions to the AusNC aim to include collections of transcribed speech with aligned audio files, as well as visual records of interaction (video). These advances have opened up all sorts of new avenues for interesting research possibilities.

Boys and girls, women and men

Figure 3.1 shows the results of a 2002 study by Sigley and Holmes on the use of *boy* and *girl* in relation to women and men. What patterns are evident from the results, and how do they relate to the treatment of women and men in society more broadly? The three corpora involved were developed at different times. How does this relate to the results?

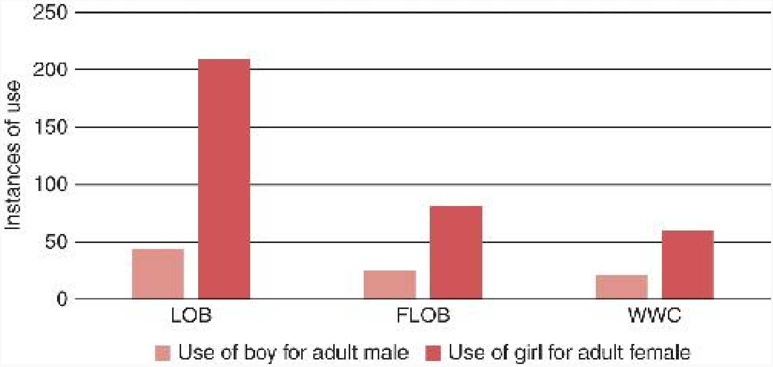


Figure 3.1 Use of *boy* and *girl* in reference to adults
Notes: There were some difficulties with assigning age classifications since these had to be inferred from the context. LOB = Lancaster-Oslo-Bergen Corpus (1961, BrE); FLOB = Freiburg-LOB Corpus of British English (1990–91); WWC = Wellington Corpus of New Zealand English (1986–90).

Observational data

Observational data are a very important foundation for linguistics research. Apart from having intrinsic value, observational data provide much of the source material for certain types of corpora described above. For example, corpora that include transcriptions of spoken interactions are based on observational data obtained from sound recordings. As the use of video recording has become cheaper and more manageable, it has increasingly been used to augment sound recordings. One important advantage of video recording over sound recording alone is that it allows the researcher to take gestures, gaze and other contextual cues into account.

A potential disadvantage of all these methods is the **observer's paradox**. People often change their behaviour if they are aware that they are being observed. How, then, is it ever possible to know what people do when they are unobserved? Since it is a basic legal and ethical requirement that people provide informed consent before participating in research, this problem is to some extent unavoidable. We will return to these vexed questions in a moment.

To understand why there is a disconnect between what people say and what they do, it is useful to think in terms of cultural models. James Paul Gee suggests that cultural models tell us what is 'normal' or 'appropriate' to expect in particular culturally defined situations. Gee identifies three different types of cultural models that can be used to show how complex the relationship between behaviour and self-reporting really is:

- *Espoused models* are those we consciously value and seek to follow. We are likely to self-report positively in relation to these. For example, we may report that we are more committed to a fair distribution of the housework than our behaviour indicates.
- *Evaluative models* are used to assess ourselves and others (consciously or unconsciously). These may not be the same as our espoused models. For example, we may 'know' that swearing is impolite but secretly admire people with flair in that area.
- *Models-in-(inter)action* guide our behaviour (consciously or unconsciously). These may be very different from the models we espouse or use in evaluating ourselves or others. For example, we may say that speeding while driving is dangerous (espoused model) and we may be morally outraged when someone overtakes us in a hazardous way on the freeway (evaluative model); nevertheless, when we consider that it is safe to do so (is it ever?), we too may speed.

Observation may involve examining people doing a certain type of activity (for example, small children playing house and emulating adult talk), or recording people having certain types of interactions (for example, leave-takings in the supermarket). Recordings of unscripted interactions from radio or television are potentially useful sources of observational data. Observation can also happen less systematically as researchers notice how language is used around them. For instance, lots of interesting data about the use of *youse* in Australian speech can be collected quite quickly if you get a group of linguists or English teachers talking about where and when they have heard it used. Obviously, this sort of anecdotal data has a different status to data that has been collected more programmatically, but it has an important role to play in allowing us to develop hypotheses about how language is being used.

Yeah-no … he's a good kid

When Kate Burridge and Margaret Florey set out to investigate the use of *yeah-no* in Australian English, they started off by casually noting examples of this expression. This provided the data for their preliminary analysis and gave an indication that they were dealing with an interesting feature that deserved further investigation. They then began to formalise their investigation and set out to tape- and video-record conversational data, drawing on a number of sources for analysis: a corpus of informal conversational data, videotaped television programs (interviews, sports programs, drama) and film.

Participant observation refers to the practice of spending time becoming a member of a particular community with the intention of developing a deeper understanding of its social practices, motivations and experiences. In linguistics, participant observation is particularly associated with research involving the description of minority languages and is associated with efforts by the researcher to become a speaker of the language. In these settings, participant observation allows the researchers to observe a wide range of situations and behaviours. These observations would typically be interpreted in consultation with knowledgeable advisers from the community.

Offering tea in the Mali community of Papua New Guinea

Participant observation provides insights into language use that are not available by any other means. For example, Mali has the verb *ter* 'refuse, decline'. To really understand how this verb works in practice, you have to be part of social exchanges. Over a long period, Tonya Stebbins learned that people may decline an offer of something that they actually want six or even eight times before they feel able to accept it. It was only after the repeated experience of failing to make successful offers and asking around about how a successful offer could be made that Tonya discovered how persistent she had to be in order to persuade someone to accept a cup of tea and a biscuit!

Case studies are useful in developing an in-depth understanding of the language behaviours or language development pathways of specific individuals or small groups of people. By focusing on a small number of people, it is possible to consider their language practices in more detail and to take a greater amount of contextual information into account. For example, a researcher interested in language delay among children diagnosed with autism might explore the language development of two or three children with different traits who are being treated in a particular therapeutic framework to see whether it works equally well for all three. Although a sample size this small does not allow for generalisations, case studies of this type can provide us with new directions to consider when using other, more statistically robust methods. Since research is cumulative, comparisons can always be made with similar case studies.

Sticking with Japanese

Masae Takeuchi was interested in researching children with parents of different linguistic backgrounds, specifically the Japanese–English combination in Australia. What puzzled her initially was why some children continue to speak Japanese with their Japanese-speaking parent while others give up. She selected four Japanese mothers and their primary school-aged children from Japanese–Australian families, and recorded their conversations over a period of time. The highly personal and individualist nature of this sort of case-study research has all sorts of advantages, and it has been widely used in investigations of childhood bilingualism, and second-language learning, teaching and use.

Interview data, questionnaires and focus groups

The use of corpora and observational methods allows us to find out how people employ language in fairly natural ways. Another area of interest in linguistics relates to people's attitudes towards different types of language. For example, there has been a lot of recent research aimed at revealing perceptions of and attitudes towards English as a global language (an uncomfortable experience for many people around the world and a controversial area).

The research techniques we cover here all involve approaching people to find out what they think about a range of language-related issues. Doing this well requires some thought on how questions are framed in order to elicit the target information, as well as how the questions will be organised in relation to each other.

An important limitation of these approaches is that they rely on self-reporting. People can report what they think much more reliably than what they do. Observational methods are more useful for issues of practice while interviews and the like give good information about attitudes. For example, one method for avoiding the observer's paradox is to make use of indirect approaches. This might involve asking a series of questions about a significant event in a person's life to elicit relatively unmonitored talk that you could use to study their pronunciation. Since this approach involves the deception of participants, it needs careful planning and justification to obtain ethics clearance.

Overcoming the observer's paradox

As we have been discussing, linguists aim to collect natural language, but their presence destroys this possibility. In fact, it is a well-known dilemma of any linguistic fieldworker. As William Labov once put it, 'Our goal is to observe the way people use language when they are not being observed'. His solution was to ask people to talk about situations in their lives when they thought to themselves, 'This is it! I'm going to die.' The emotional responses of these near-death experiences triggered the 'forgotten recorder effect', and the researcher got naturalistic speech. Not everyone has found this a useful question, however. Crawford Feagin tried this approach in Alabama, and most of the people simply responded to the opening question (something like, 'Have you ever thought you were going to die?') with 'No.' Sociolinguists Janet Holmes (in New Zealand) and Peter Trudgill (in England) experienced the same response to the 'danger of death' question in their interviews.

Clearly, questions need to be tailored to local conditions and speakers' tastes. Crawford Feagin writes that, 'In Anniston, after the danger-of-death question proved unsuccessful, I discovered that the question "Have you ever heard of anybody seeing a ghost around here?" often elicited long, elaborate narratives of local mayhem and murder from older working-class speakers.'

What questions do you think might elicit passionate speech from your friends and family? Bear in mind that something people would look back on with emotion might well involve traumatic experiences that participants may not want to revisit.

The size of the sample continues to be an important aspect of these study designs. As with the other approaches we have covered so far, a smaller sample size relative to the population will obviously have more limited application than a large sample; however, there are many reasons why a small sample size may be an appropriate choice in practical terms. Often a great deal of thought is devoted to the identity of the people who are invited to participate in the study. They need to be representative of the target group, however that is defined, and also reliable and cooperative.

Interview data involve the collection of answers to a series of questions that have been planned in advance. The interviewer works with participants to collect their answers to each question. Interviews can be structured, semi-structured or unstructured. These different categories reflect different degrees of reliance on the list of questions. A less-structured interview will have a planned list of questions but will be administered in a more conversational way, with the interviewer keeping track of which topics have been covered. At natural breaks in the flow of the conversation, the interviewer can prompt the participant for responses to topics that have not yet been covered. Depending on the nature of the basic research question, a more structured approach may sometimes be useful – perhaps because it ensures that all the participants in the research are treated in a similar way.

Focus groups involve discussion within a small group of people on a specific topic. They are useful for finding out about what people think because they provide a forum in which people can share their views and ideas. The trick with focus groups is to find a way to get everyone focused on the intended topic, but also leave enough freedom in the group so that people get to explore the issues involved. In some ways, running a focus group is a bit like running an unstructured interview, since the role of the facilitator is to keep things on track and prompt for responses on areas that have not been covered.

Gerontolescence and the longevity revolution

A team of linguists at Monash University has been investigating the sort of language that Australians use to talk about (or not talk about) getting old. Prolonged old age has become one of the most significant medical and societal breakthroughs of our time, yet ageing is a topic that remains well and truly a taboo within our society. This is reflected in the behaviour of words that are associated with ageing. The word *old* has shed its positive associations of wisdom and respect, and once very positive words like *senile* have been quick to deteriorate in their connotations. Even the label *elderly* now carries negative undertones.

So, these linguists have gathered focus groups of older Australians from a wide range of occupations and life experiences, and have organised individual interviews with 'successful agers' and those who might be thought of as role models for society. The sorts of questions they have been asking include: How do you think people talk about ageing? What is your experience of contemporary views of ageing in Australia? What do you think it means 'to age well' and how would you define 'successful ageing'? What terms for 'older Australians' are you familiar with? Who is referred to as *old* or *elderly*, or as a *senior citizen* or *senior*? What do you think of novel terms such as *gerontolescence* (a blend of gerentology + adolescence) or *the chronologically gifted*? and 'How do you prefer to be described?' Recent research has found that 60-to 70-year-olds are the happiest people alive. What would you say to this?

By exploring these ideas, the team has stated that it hopes to provide the vital linguistic provisions needed to help us all to grow old gracefully and 'successfully'.

Questionnaires are a useful way to collect data from a larger range of participants, since they are typically completed by participants independently of the researcher. If they are carefully planned, they will be easier to interpret than a series of one-on-one or focus group interviews, and it is more likely that you will be able to make reliable comparisons across the results.

One common challenge with questionnaires is that there is often a low rate of return. This is because it can be difficult to motivate people to participate in responding to a questionnaire, particularly if they either do not understand or do not care about the topic. The trick, of course, is to find a topic about which everyone has an opinion. Monash University researcher Allie Severin found exactly that topic!

'Zombie rule that I don't live by': Measuring language attitudes with a changing yardstick

In 2013, Allie Severin conducted research on the linguistic niggles of Australian speakers, trying to find out how much they cared about matters of usage. She was particularly interested in the opinions of younger speakers. (It is, after all, a well-trotted-out trope that young people do not care much about language use and that language standards are deteriorating!) Severin had an extraordinary response to her call for participants – almost 400 people completed her online survey, willingly expressing their opinions on matters of English usage often identified as prescriptive bugbears.

Severin's findings were interesting. Here is what she wrote for a Boobook Education blog.

> It turned out that younger and older speakers alike believed in the necessity of adhering to Standard Australian English. Regardless of age, people were unforgiving when they came across one of their linguistic 'pet peeves'. For example, in response to the phrase *I don't know nothing*, a younger participant declared, 'I think people are stupid when they use this' – while an older one stated, 'This just sounds very lower class or uneducated.' Yet there were still differences in the ways that age groups approached the matters of usage they disliked. For the most part, older speakers refused to accept the use of non-standard or supposedly 'incorrect' English in any context. Younger speakers, meanwhile, demonstrated awareness of alternative linguistic forms but remained critical of them. In the cases where younger participants viewed a form as non-standard (rather than simply incorrect), they were more willing to accept its use within certain contexts. They may have disliked double negatives, but the majority still recognised that *I don't know nothing* is the norm for speakers of some English dialects. In effect, younger participants were more likely to consider the social and cultural background of speakers when deciding whether a form was acceptable or not.
>
> However, when it came to things that they considered to be incorrect, they were just as critical as their older counterparts. It didn't matter who you were, if you spelt *you're* as *your*, they weren't going to be happy with you. But there were interesting differences in what provoked this reaction. Older speakers, for example, cared a lot about use of the word *gotten*; younger speakers were almost entirely unaffected by it. What united both groups was the force of their emotion – when confronted by something they didn't like, they were all incredibly passionate.

By the way, 'Zombie rule that I don't live by' was the description offered by one of Severin's participants for a prescriptive rule of English no longer relevant to Australian English speakers.

As with the other methods discussed, there is a relationship between how many responses are collected, and how reliable and generalisable the findings will be. Another challenge is that participants may have trouble interpreting particular questions. If they are responding to a written questionnaire, there is no chance for them to ask clarifying questions, and they may miss the point of the question and answer inappropriately, omit the question or, worse, abandon the survey altogether if it is too difficult to follow. Even though Severin constructed her questions carefully so participants would focus on the feature being investigated, sometimes they would react to completely different aspects of the sentence. For example, despite being asked to comment on grammatical features, some participants based their responses on semantic content. For instance, six labelled the sentence *I like jam, vegemite, and peanut butter* either ‘acceptable when used by others’ or ‘unacceptable’ and then wrote in the comment section that they did not like these foodstuffs – Severin was interested in their views on the punctuation, not their food preferences! A small number of respondents reacted to non-target forms in sentences; for example, in the sentence *It’s simple: your just plain wrong*, three based their choice on the colon use, rather than the misspelling of *you’re*. In the sentence *Apple’s for sale*, seven respondents based their decision on the possibility that the sentence referred to Apple Inc. rather than the plural noun *apples*, which was the intention. It is often impossible to anticipate how participants will respond to survey questions!

Experimental data

Another way to discover what people actually do (as opposed to what they say they do) is to run an experiment. In contrast to observation, which involves reports of fairly open-ended and unconstrained behaviour, experimentation involves the creation of highly controlled tasks that participants are asked to complete. Sample sizes are likely to be relatively small, since participation in an experiment is likely to involve greater time and effort than the completion of a survey (for example, an experiment may require participants to go to a particular place in order to complete tasks in a particular environment or with particular toys or tools). **Experimental data** generally are interpreted in a quantitative framework (analysed using mathematically based methods, such as statistics). This means that a knowledge of statistical methods is useful for conducting and understanding this type of research.

Do swearers lead less stressful lives?

It is a common observation by some in the community that those who condemn swearing are ‘uptight’. This opinion is now buttressed by new experimental research conducted by England’s Keele University. The study ‘Swearing as Emotional Language’ was carried out by psychologist Dr Richard Stephens and presented at the 2014 British Psychological Society annual conference in Birmingham.

Dr Stephens and his team were the ones who a few years back famously conducted the memorable ‘Ice Water’ experiment, showing convincingly that swearing helps us to endure pain. Participants were asked to plunge their hand into a tub of ice water and hold it there for as long as possible while repeating a single swear word of their choice. They then repeated the experiment, but this time they were only able to say an everyday word, one that could be used to describe a table (such as *wooden* or *round*). (How did they get this past the ethics committee, we are wondering?) The majority of participants were able to keep their hands in the ice water for a considerably longer period of time when using the swear word, supporting the link between swearing and a higher pain threshold (although this effect is diminished for habitual swearers, and Stephen Fry’s re-enactment of the experiment on the television program *Planet Word* confirmed this).

The results of the more recent study ‘Swearing as Emotional Language’ showed that raising people’s emotional arousal level makes them more proficient at swearing. Participants were asked to play either an aggressive video game or a golf video game. After their sessions, those playing *Shoot-em Up* could recall significantly more and a wider variety of swear words than those who played golf (not a surprising finding, to be sure, though one that doesn’t quite mesh with one of the current authors’ experience on the golf course as a caddy many years ago). Dr Richards told a journalist who reported on the research, ‘As we looked into swearing further, it became apparent that it’s actually emotional language, and can make you feel better in certain situations.’

So swearing is closely related to emotion and makes us feel better – not exactly groundbreaking research. As we will explore further in Chapter 16, many studies, carried out over many years, have revealed the clear sociocultural and psychological benefits of using taboo words and phrases. But, it is nonetheless interesting to have additional back-up from empirical research, that swearing is an emotion-based coping mechanism that makes us feel more resilient – a further confirmation that ‘bad’ language isn’t always bad.

3.2 Uses of linguistic data

Since people talk and write all day, every day, and since many people have strong views on language and an interest in supporting language learning in the community, language data can be surprisingly easy to collect once you know how. Collecting useable data relevant to specific research questions takes a bit more skill and planning, since research then needs to be developed with an eye on what kind of data will be relevant and how such data can be used. In this section, we outline the basic approaches to using data about language in research.

Researchers often draw a broad distinction between quantitative and qualitative methods of analysis. This section provides a brief overview of these two approaches, then relates them to different types of work in linguistics. Although quantitative and qualitative methods of analysis are more frequently used in some types of research than others, it is worth bearing in mind that many types of research involve a 'mixed-methods' approach, in which a range of different types of data and methods of analysis are brought together to provide more meaningful answers to particular research questions. For example, a description of an endangered language is likely to include simple statistics about the size of the speaker population along with analysis of the domains in which the language is used, in addition to materials that describe the language itself.

Quantitative analysis

Quantitative methods of analysis involve answers to questions about *how much* or *how often* something occurs and *how similar* or *how different* one thing is from another. This is useful as a means of answering questions within existing theoretical frameworks. For example, accommodation theory (covered in Chapter 12) leads us to expect that people will accommodate their speech to others (for example, by switching languages) when the rewards of establishing and maintaining social relationships outweigh the costs to a person's identity of doing so. Research that supports this theory measured how much participants were prepared to help someone whose language was different from their own. Researchers were able to show that the more the person asking for help matched the language of the participant, the more willing the participant was to help. Other research in this area considered whether the different social status of the languages concerned was significant, and measured whether there was a correspondence between language status and frequency of accommodation.

Quantitative approaches are extremely helpful in allowing researchers to manage large amounts of data and draw meaningful results from the information they collect, provided that a suitable research question has been posed and relevant, well-organised data have been collected within a coherent study design. Such approaches are used across many sub-fields of linguistics, from counts of text frequency for particular words in historical linguistics to the use of standardised tests to measure language development in speech pathology. Quantitative analysis is also used to search for patterns in corpora for cross-linguistic data (the findings on *bloody* reported on earlier are a simple example of this).

Qualitative analysis

In contrast to quantitative analysis, **qualitative** analysis is concerned with how things happen and what they might mean. Rather than counting things, this approach is concerned with interpreting them. Whereas the data collected for quantitative analysis must be relatively uniform in order to be counted, the data used in qualitative analysis may be quite diverse, since meaning-making can benefit from many different perspectives on a research question. This can result in large volumes of apparently unrelated data. For example, research on language use among speakers of multiple languages could involve a range of observational data (interviews, focus groups, and observations in planned and naturalistic settings).

One important strategy for managing disparate data is to find ways to keep the data organised. Coding the data allows the researcher to categorise them in different ways. For example, research on multilingualism might code the data for the location and level of formality of each interaction, as well as the dominant language, language attitudes, age, and gender of each speaker. It is useful to note that the types of content that may be significant include both the manifest things that are said and shown on the surface, and latent meanings that reflect the researcher's understanding of the meaning of the material. Thus, when asked directly about language attitudes, participants will make statements about their views, but in other contexts the same or different views may also be inferred by their behaviour.

Qualitative analysis is relevant to a broad range of areas of inquiry within linguistics, from evaluations of community attitudes towards various languages to exploring how teachers correct students' language errors in an English as a second language (ESL) classroom, and the types of impacts these actions might have.

Learning about texting

One study from 2003, by Crispin Thurlow, explored the development of texting behaviours and language forms involved in a collection of text messages from a cohort of university students, creating a small corpus of messages. The findings mix qualitative and quantitative methods of analysis. The average text length was fourteen words. The purposes of the messages were categorised as shown below. The percentage refers to the frequency of each message type:

- *informational-practical orientation (*straightforward requests for information): 14 per cent
- *informational-relational orientation* (solidarity information exchanges or requests for personal favours): 8 per cent
- *practical arrangement orientation* (plans to meet or the coordination of household expeditions): 15 per cent
- *social arrangement orientation* (recreational planning and other social arrangements): 9 per cent
- salutary orientation (usually very brief and often flippant; simple, friendly greetings): 17 per cent
- friendship maintenance orientation (apologies, words of support and thanks): 23 per cent
- romantic orientation (expressions of love, intimacy and affection): 9 per cent
- sexual orientation (explicit sexual overtones): 3 per cent
- chain messages (epigrams, jokes or wordplays): 2 per cent.

It is interesting to note how the messages are spread across such a range of communicative types, but also how focused the majority of the message types are on maintaining social connections.

The last few decades have seen a surge in quantitative approaches within linguistics. In a 2018 colloquium on methodologies across the humanities and social sciences, linguist Bernd Kortmann attributed this quantitative turn to the availability of bigger and better electronic corpora, as well as more sophisticated statistical tools able to handle large and complex linguistic data. Another contributing factor is the increasing importance of psycholinguistic and neurolinguistic experimentation in language processing and language variation research (something we explore in later chapters). Linguistics has always crossed traditional discipline boundaries, and is a field in which both the quantitative and qualitative methods appear to be playing nicely alongside each other as ‘a healthy equilibrium, if not mutual reinforcement, of quantitative and qualitative approaches’, to quote Kortmann. This productive relationship, according to Kortmann, places linguistics in the position of being able to forge new alliances between the STEM (science, technology, engineering and maths) disciplines on the one hand, and the HASS (humanities and social sciences) disciplines on the other.

3.3 Ethics in linguistics

The final section of this chapter is concerned with the relationship between linguists and the people they research. Human research ethics has its antecedents in the field of medicine. The systems that many countries use to protect people and ensure they are treated ethically when involved in research are founded on a medico-scientific model that has been extended to encompass other types of research as well. We begin with an overview of the key concepts that organise thinking about ethics in the medico-scientific model before turning to some broader interpretive issues that are becoming increasingly significant for social science researchers.

Contested values in language research

Knowing about language and understanding the ways in which wellbeing and opportunity can be limited for speakers of non-prestige varieties pose professional and personal issues for linguists. John R. Rickford describes a foray into advocating on behalf of African American English speakers for educational reforms that supported the use of African American English in schools. In what became known as the ‘Oakland Ebonics Controversy’, the Oakland (California) school board passed a resolution on 18 December 1996, recognising the legitimacy of ‘Ebonics’ (or what linguists label African American English or AAE). The fury of media criticism that erupted highlighted the popular misconceptions around vernaculars and the difficulties of getting the wider community to accept even the idea of non-standard varieties as legitimate linguistic systems, and not just blots on the linguistic landscape that needed to be eradicated. Rickford worried about whether his colleagues would be willing to engage with this issue.

> As it turns out, I need not have worried. To my amazement, syntacticians, semanticists, sociolinguists and linguists of other stripes and persuasions rose to speak in support of the proposals – if anything, they wanted to make them even more radical – and a four-point resolution was unanimously approved. Among other things, it affirmed the 'rule governed and systematic nature of Ebonics', and pronounced the Oakland school board's decision to take it into account in teaching Standard English 'linguistically and pedagogically sound'.

Human research ethics

Human research ethics centre around three core sets of concepts: risk, benefit and consent. In essence, in determining whether or not research is ethical, the researcher and any committees or reviewers who are required by the state to monitor research must weigh the balance of **risks** and **benefits** associated with the work to ensure it is worthwhile. In ethical research, the potential benefits outweigh the risks. Participants in the research need to understand the risks involved and the potential benefits to themselves or the wider community. There is also a clear requirement that participants should not be involved unless they have given their informed **consent**. This means that they understand and freely agree to participate in the research.

In weighing up the balance of risks and benefits, a number of other factors become relevant. One of the most important is to consider whether or not the research has wider **merit** – is it worth doing? Good research is generally expected to make a contribution to knowledge that others can recognise as valid, and that is shared honestly and openly. **Justice** is also an important value. Ethical research is just when it does not involve exploitation or undue hardship for participants, and ensures that access to the benefits from the research is fair and equal. Finally, **respect** for the individual includes researchers taking into account the values and beliefs of the communities in which they work, and protecting their privacy, both as individuals and as communities.

The values of justice and respect are uncontroversial for most people. It also makes sense that research is not worthwhile unless the benefits outweigh the risks, and that people should be informed and not coerced or tricked into participating. Nevertheless, ensuring that research is ethical can sometimes be challenging in practice. For example, there are many contexts in which the researcher knows far more about a topic and has a much deeper understanding of the likely significance or potential impact of the research than the participants. It is also sometimes difficult to really know what benefits are likely to result from a research project or what kinds of protections are most appropriate in particular communities.

We noted earlier in this chapter that the observer's paradox refers to the problem that informed consent prevents the researcher from ever really obtaining naturalistic data, since the situation without the effects of the observer can never be recorded. In terms of collecting linguistic data, one common strategy is to provide participants with general descriptions of the research that do not flag the issue being studied. Under special circumstances, it is also possible to get approval to conduct research that involves deception of participants, provided the research has sufficient potential benefit and there is no alternative means of obtaining the data.

Recording people in conversation is a typical scenario in which the observer's paradox applies. In this situation, it can be difficult to obtain naturalistic data because people typically become self-conscious when they know they are being recorded. It might seem that the obvious thing to do would be to conceal the recorder and make a recording without the consent of the participants. Not only would this be unethical, but in many places it would also be illegal. Fortunately, there are a number of things that researchers can do to minimise the effects of self-conscious participants when recording. Tape recorders cannot be concealed, but they can be placed unobtrusively. Naturalistic data often starts to emerge as the conversation warms up, so providing compelling topics for discussion and making use of sections of talk later in the recording are also useful strategies. Nonetheless, the forgotten tape recorder does come with its own set of ethical issues – is a forgotten device really so very different from a hidden device?

Isabelle Burke is a Monash University researcher who investigated the current state of Australian English syntax using a slightly unorthodox corpus: the reality television show *Big Brother*.

Big Brother is watching (your syntax)

Bear in mind that you can legally record public talk, and that broadcasting media, such as television and radio, offer hundreds of opportunities to capture different kinds of (fairly) naturalistic data (think of talk shows, panel discussions, debates, political discussions and so on).

For those unfamiliar with the show, which ran for several seasons, and was a global franchise, *Big Brother* involved secluding a number of housemates in the *Big Brother* house, where their every movement – and more importantly, conversations – were recorded 24 hours a day for several months. Certainly, the housemates were aware that their talk was being recorded at all times and they did show their awareness of this situation by occasionally referring to the recording process, or that they were being overheard. But, because they had unscripted and unplanned interactions with one another 24 hours a day for months on end, it is hard to imagine they would have maintained a constant attentiveness to any audience. Burke concluded that the data that emerged from *Big Brother* were 'neither completely authentic, nor artificial, but something in between'.

Without doubt, these ongoing recorded conversations are an excellent vehicle for examining casually spoken Australian English and the grammar changes within that variety; and they are a particularly welcome data source because grammatical research requires such large volumes of language – far more than is needed for studies of other features, such as pronunciation or word structure.

Power, representation and authorship

One of the risks to participants in many kinds of social science research is that they are also the subjects of the research. This means that the researchers control how participants are portrayed in their findings. On a community level, this involves such issues as the ways people are defined in relationship to one another. Some of these issues can be difficult for the researcher themselves to spot. For example, the question 'Do women speak differently from men?' may seem innocuous enough at first glance; however, it involves an underlying assumption that men's speech is 'normal' while women's speech is at least potentially 'deviant' – a long discussion on the topic can only serve to reinforce these underlying propositions. Furthermore, the fact that these portrayals of men and women are not stated makes them much harder to identify and to question.

On an individual level, a standard approach to the problem of unwittingly unfair portrayals of participants in the medico-scientific model is to make participants anonymous. This approach works quite well in certain types of social science and linguistics projects, particularly those with very large numbers of participants representing broad cross-sections of a society. However, within linguistics there are some situations in which this approach is unacceptable to participants, and others in which it is simply impossible to ensure anonymity. Like consent, the identification of participants within research sometimes needs to be discussed and negotiated until participants are satisfied that arrangements will be fair and will minimise the risk of harm; for example, both these situations arise in research on small endangered languages.

In very small communities, people are well known to each other and the total pool of participants is inevitably limited. Under these conditions, it is likely that other members of the community will be able to identify a speaker based on the content of their speech (such as in a quote to illustrate a certain grammatical construction). Speakers themselves also need to take care that the things they discuss do not include material that may upset others in their community. Since matters of controversy are culturally laden, it can be very difficult for the researcher to identify contentious material out of context without advice from a participating member of the community.

The other problem with anonymity is that by reducing the visibility of participants in the research, their contributions are also diminished. For example, many Aboriginal people in Australia deeply regret that there is not more information about the people who contributed to archival materials relating to their languages. For many Aboriginal people, knowing who gave language materials to researchers, and especially being able to trace family relationships to these people, provides a critical feeling of authenticity in relation to the materials. Unfortunately, the materials are catalogued under the names of the recorders, who were often surveyors or missionaries, which places a barrier between the community and the words of their ancestors. When texts are collected in these communities today, people are understandably actively concerned to protect their intellectual property and identify themselves as authors of their own work within a project.

Relationships and research partnerships

Researchers within linguistics have long recognised that it is also possible to seek to increase the benefits of the research that flow back into the communities where they work. Cameron and colleagues describe a series of different stances that the researcher can take in relation to managing relationships with participants, including the acknowledgement of risks and the identification of benefits in linguistics research.

Ethical research assumes that the researcher does research ‘on’ participants. This approach to research involves the extension of the medico-scientific model into the social sciences, as described above. Participants are to be protected from undue harm, benefits should flow back to them and they should generally be able to exercise informed consent. This is the full extent of the researcher’s obligations. The issues just flagged in relation to how participants are presented, while they might be raised in conjunction with the findings, do not inform the progress of the research itself. In this model, the question ‘Do women speak differently from men?’ would be acceptable provided no one was unduly inconvenienced in the search for evidence and there was some sort of benefit to someone.

Advocacy research is both ‘on’ and ‘for’ the participants in the research. This approach involves researchers using their expertise and access to resources to pursue research or to advocate publicly in the interests of the communities where they work. This may involve serving as expert witnesses, writing letters of support, accumulating compelling evidence on significant issues and raising awareness in the wider community about issues of concern. The difficulty with this model is that it does little to disrupt the underlying assumptions that researchers bring to the areas they choose to study. In this model, the problematic framing of the question ‘Do women speak differently from men?’ might be contested by the researcher on behalf of all women, but the upshot of the whole debate will remain within academic circles and be determined through established means. The question itself is unlikely to change.

Empowering research involves research that is ‘on’ and ‘for’ and also ‘with’ participants. In this approach, the power that naturally accrues to researchers as those who ask the questions is defrayed among participants by allowing them to ask questions back and to raise their own topics for inquiry. In this model, the question ‘Do women speak differently from men?’ is more likely to change since women participants are likely to offer their own questions about language and gender, and to trouble the assumption that they are necessarily the ‘deviant’ party in the equation (for example, by pointing out that their language is stereotypically said to be more polite and closer to the standard than the language of men). Cameron and colleagues argue that there is potential for better research in this model, since the process of working with participants in this more equitable way may lead to unexpected insights. In connection to our hypothetical example, for instance, it might lead to some other research questions about the patterns and norms of both men's and women's speech that give legitimacy to each (if indeed differences were found to exist).

Further reading

A wonderful book full of useful information to help you 'do linguistics' is Bauer (2007); in it you will find the fundamentals of linguistics outlined, as well as tips on such things as how to write assignments and make sense of linguistic notation. As the preface says, it 'helps the student of linguistics with all the things that the instructor doesn't quite get round to'.

Wray and Bloomer (2013) provide more detail about the different approaches to research in linguistics; this book is a useful next point of reference for students. It also contains a great deal of useful advice about designing small research projects and writing them up. A more advanced but very practical text that describes research methods in linguistics is Litosseliti (2010). The text by Heigham and Croker (2009) has a particular focus on qualitative methods in relevant areas of study while Rasinger (2008) focuses on quantitative methods. A useful resource on the role of quantitative and qualitative methods across the humanities and social sciences is the 2018 Freiburg Institute for Advanced Studies lecture series, where you can find the podcast of Kortmann's lecture (on linguistics).

Musgrave (2012) provides an excellent account of what was involved in establishing the Australian National Corpus (AusNC); Rodriguez Louro (2013) details the research on the 'quotative *like*' down under; Musgrave and Burridge (2014) draws on corpus linguistic data to investigate the development and cultural significance of *bastard* and *bugger* in Australian English vernacular. Labov (1972c) is the now famous account of linguistic methodology and the quotation cited in this chapter appears on page 61 of Labov's work; the suggestions for eliciting naturalistic speech come from Feagin (2004), and the cited quotation appears on page 30.

Ochs (1979) raises a number of issues relating to transcription that deserve attention in any project that involves the rendering of recorded speech as text. The examples from accommodation theory are drawn from papers by Simard, Taylor and Giles (1976) (on willingness to help), and Aboud (1976) (on accommodation and language status). The material on cultural models from Gee (1990: 83) provides one framework for understanding the ways in which our ideal selves do not always relate directly to our actual behaviours.

Stephens (2013) presents a summary of recent research on the psychological functions of swearing. Takeuchi (2006) describes the research we mention about raising bilingual Japanese–English speaking children in Australia, and Thurlow (2003) provides an interesting review of common uses of text messages; Severin's (2017) study focuses on online normative language debates on the discussion forum website Reddit (specifically the subreddit R/badlinguistics, dedicated to the sharing of 'all the worst linguistics-related content on Reddit and the Internet').

Rickford (1999) provides an account of his linguistic activism. Finally, Cameron et al. (1992) presents a well-considered discussion of the different relationships that linguistics researchers can have with the communities and individuals with whom they work.

The Australian Linguistic Society (ALS) (the principal professional body concerned with the discipline of linguistics in Australia) has drawn up a statement of ethics that applies to its members, and to persons directly employed by members, in the conduct of linguistic research. You can find the list of ethical principles at the ALS Policies web page (https://als.asn.au/AboutALS/Policies).

Exercises and discussion points

1 Approaches to language

For each of the following research questions, would you use a quantitative, qualitative or mixed-methods approach? What types of data collection would be required?

(a) You want to know whether being an only child has an impact on the language learning process.

(b) You want to know whether newspaper (and/or online) job advertisements make assumptions about gender.

(c) You want to compare the salutations in emails with those in letters and phone calls.

(d) You want to compare a real conversation with a scripted conversation from a television program.

2 Corpus activity

Compare the frequency of modifiers describing appearance and behaviour with the words *boy* and *girl* in the 2002 study by Sigley and Holmes (see Figure 3.2). How would you explain their different treatment? To what extent do you think this reflects cultural models associated with particular genders?

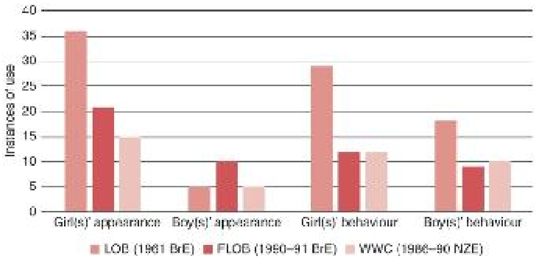


Figure 3.2 Appearance and behaviour in references to *boy(s)* and *girl(s)*

Notes: There were some difficulties with assigning age classifications since these had to be inferred from the context.

LOB = Lancaster-Oslo-Bergen Corpus (1961, BrE); FLOB = Freiburg-LOB Corpus of British English (1990–91); WWC = Wellington Corpus of New Zealand English (1986–90).

In relation to this research, the authors have noted the following important limitations.

- The gender of the writers of each text was taken into account in the design of the corpora, although it was unknown in many cases.
- Writers sometimes use techniques like parody to critique social and cultural norms. It is not possible in this kind of research to ascertain how many uses of *girl* were intentionally inappropriate for this type of purpose.
- The corpora involved are relatively small. In a larger corpus, the preferences of individual authors would be less noticeable than they probably are here.

What other things would need to be taken into account in interpreting these findings?

3 Focus on compliments

(a) What are the cultural norms around giving and receiving compliments?

(b) When and how are compliments usually given?

(c) What is your usual response to a compliment? What do you say? How do you feel? Why?

(d) What would prompt you to compliment someone else?

(e) When you compliment someone else, what do you say? How do you feel? Why?

4 A corpus investigation

As described in this chapter, the AusNC offers all sorts of possibilities for the exploration of spoken and written Australian English, even opening windows on historical variation in the early history of the language. This exercise is designed to familiarise you with the various collections on AusNC (http://www.ausnc.org.au). Visit the website and respond to the following.

(a) Select one of the collections and choose an aspect of a linguistic feature about which you are curious (for example, you might want to know about the functions of a discourse marker such as *yeah-no*, or perhaps you want to know whether females use 'emphatic *like*' more than males, as in *It's like so excellent*).

(b) Outline the advantages of a corpus-based investigation of language (using AusNC) as opposed to one that relies on the linguist's introspection, intuition and anecdotes.

(c) What aspects of language can be studied using a corpus?

(d) Can you see any drawbacks to a corpus investigation?

5 Ethical, advocacy or empowering research

Consider the following scenarios and evaluate them in terms of the concepts of ethical, advocacy and empowering research.

Scenario 1

A PhD student travels to a remote community on a Pacific island to prepare a grammar for her dissertation project. On arriving in the community, she is welcomed warmly and told that people would like her to record as much of their language as possible. However, she struggles to find anyone who is willing to spend time working with her on the tasks of recording, transcribing and translating texts. Without a corpus, it is difficult for her to see how she can complete her project.

(a) What type of approach do you think the student is bringing to the project?

(b) What should the student do?

(c) What kinds of barriers might there be to greater involvement from the community?

Scenario 2

A research project on Turung was partly funded by a grant from the Endangered Languages Documentation Programme, based at the University of London. A focus of the project was the methods used to collect, analyse and store data, involving five stages:

- recording data
- capturing the recordings digitally
- analysing the data (this involved transcription, translation, glossing and annotation)
- archiving the data
- mobilising the data (making it available to other researchers and to the source community).

(a) What types of resources do you think would be required in each of these stages?

(b) What types of support would be required from the community?

(c) If you were a fluent first-language speaker of the language being documented, would your answers change?

Research project

Collect ten recipes for one of your favourite dishes (chocolate cake, for example) from a range of different sources, such as recipe books, websites, specialist magazines and so on; compare them with each other. Most recipes have at least two different text types: a list of ingredients, and a set of instructions about how the ingredients, should be prepared.

1 Do all the recipes you've collected share the two features mentioned?

2 Are there any other types of text included in any of the recipes?

3 What other differences can you find between the recipes? (For example, you can check whether the ingredients lists are ordered differently in different recipes.)

4 Do all the recipes use similar styles of language?

5 If there are differences between the recipes, what do you think has motivated them?

For more activities, please go to www.cambridge.edu.au/academic/linguistics.

Part 2

‘In the beginning was the word’

Words, word structure and meaning

4

What's in a word?

◆

> I know words. I have the best words.
>
> Donald Trump, 30 December 2015

4.0 Introduction

The word is probably the most familiar linguistic concept to non-linguists. This is because the lexical aspects of languages are generally more accessible than other aspects. People love playing with words and arguing about them, but most of all they take great delight in inventing new ones (Oxford Dictionaries Online adds hundreds of new expressions, or **neologisms**, to its database each month). Neologasm, a term captured by the *Urban Dictionary*, says it all: 'the intensely pleasurable sensations generated by using, hearing or coining a new word or phrase (that doesn't suck)'. Leximania, which is commonly defined as 'a compulsive desire to invent new words' starts in childhood and stays with us as we grow up.

People are also fascinated by curious facts to do with the ins and outs of the **lexicon**, that is the vocabulary of their language. English speakers are always asking us intriguing things such as: What is the longest word in the language? What word rhymes with *orange*? Is there a word to describe someone who drinks their own bathwater? How many words do speakers know? and What are the three words that end in *-gry*? (Although, since *hangry* 'irritable because of hunger' was added to the dictionary in 2018, this last question is not such a challenge.) There are websites devoted to the 'Word of the Year', favourite words, least favourite words, dead words, new words, peculiar words, clichés and many other topics, all revealing people's fascination with vocabulary.

So, this chapter is all about words. It starts by addressing the status of the word within linguistic analysis, and attempts to pin down a definition. We then look at some of the characteristics of words and their power. We explore the reasons why words disappear, but mostly we look at the ways in which we create new expressions.

4.1 Defining a word

> Definitions are like steps cut in a steep slope of ice, or shells thrown on to a greasy pavement; they give us foothold, and enable us to advance.
>
> Samuel Butler (1890), 'Thought and Language' lecture

Does the **word** form a natural unit for speakers? If so, what are the criteria for deciding whether or not a string of sounds constitutes a word? The first question is easy to answer. As speakers of English, you will all recognise that the sentence with which we started this chapter has eleven words (ignoring the hyphen for now):

(4.1) The word is probably the most familiar linguistic concept to nonlinguists.

Hopefully, you all agree that the above sentence has the correct word division and that what we now write would be completely wrong-headed:

(4.2) Thew ordis pro bably themost fam iliarlinguist icconceptto nonling uists.

You make the appropriate word boundaries, even though the sentence could be written as a continuous sequence of sounds. After all, speech is seamless, and in normal rapid speech one word runs into the next:

(4.3)
Thewordisprobablythemostfamiliarlinguisticconcepttononlinguists.

As psycholinguist Steven Pinker explains, what we do is ‘hallucinate’ a word boundary whenever we reach the end of a stretch of sound that matches up with an entry in our mental dictionaries. Although the stream of sound can sometimes coincide with more than one entry (so one person’s *euthanasia* might be another’s *youth in Asia*), generally we seem to be hallucinating word boundaries in the same way. It can’t simply be that we know how to read and write, and are therefore familiar with writing conventions that place little white spaces between certain sequences of sounds: speakers of unwritten languages have similar intuitions when it comes to segmenting utterances into word units. So, there is something natural about the word; however, actually defining *word* is not easy. There have been many attempts to pin down descriptions that accord with speakers’ intuitions, and it becomes even more tricky when we start looking at words across languages – in fact, entire books and conferences have been devoted to the topic of ‘what is a word?’

Words have different aspects, which can be helpful for establishing wordhood. We will look at three of these.

Three facets of wordhood

First, there are the semantic aspects. We could say that a word constitutes ‘a unit of meaning’. At first blush, this seems to work quite well for English, but not so well for a language like Mali. Take the word for ‘teacher’, *sunasvemka*. It breaks down into a number of meaningful parts that combine together to give the meaning ‘one who teaches/learns by trying things out themselves’.

(4.4)

su = try

nas = self (reflexive)

vem = inside

ka = masculine singular

Even in English, there are problems with the ‘unit of meaning’ definition. What do we do about the word *non-linguist* in the earlier sentence? This word is made up of two different units of meaning (*non* + *linguist*). Similarly, according to this definition, *un* would have to be a word, since *unfamiliar* is no different from the string *not familiar*. Even the string *the most familiar linguistic concept to non-linguists* might qualify as a word according to this account, since this entire phrase could be said to have a specific meaning (and to show this, we can replace the entire phrase with the single word *it*). Obviously, semantics has an important role to play in defining words but, on its own, it is insufficient.

Second, there are the phonological aspects. Words might also be thought of as units of sound structure and we might invoke phonological criteria to help us define words. Think about how you emphasise the words in *black board* versus *blackboard* – the difference in stress tells you whether you have two words or one. Some languages have regular stress patterning that is assigned on the basis of word structure. In Old English, most words carried the stress on the first syllable, as was typical of early Germanic languages.

Mali has a basic pattern as well (also reported for many Oceanic languages), whereby stress tends to be on the penultimate syllable (in other words, the second-last syllable). In many languages, though, including Modern English and Russian, stress is unpredictable. But even where stress is predictable the rules are not infallible and there are generally exceptions, such as: changes in stress patterns (*invalid* stressed differently gives us two different words, ‘illegitimate’ and ‘patient’). There are also words (or what native speakers recognise as words) that never receive stress (*a* and *the* are typically unstressed). Stress is important, but it is not enough on its own to identify words.

Other phonological properties have to do with the possible sounds and sequences of sounds that are allowed. Some languages permit a limited sub-set of consonants in word-initial position; others show restrictions in final position. English allows the cluster of sounds [kn] at the end of words, as in *beckon* [bekn], but not at the beginning (hence the silent <k> in *knee*); the sound we write as ‘ng’ (in phonetics [ŋ]) doesn’t appear initially either, although borrowings from Aboriginal languages (*ngowa* ‘mallee fowl’) and Māori (*ngaio* ‘an evergreen tree’), and names like *Ng* and *Nguyen*, are changing this. As we will see in Chapter 8, there are also pronunciation processes that occur in fluent speech and sometimes these can be useful for identifying boundaries. In Mali, for example, there are certain consonant and vowel changes that take place at the edges of words, so these become one of the defining features of Mali words.

The possibility of pausing is another way of establishing wordhood. Speakers might stop to think or correct themselves, and they typically do this between word boundaries (this can involve silences, or more usually hesitation noises of the *er/um* variety). In Mali, too, pausing takes place between words. Of course, we're talking about potential pauses and hesitations here – speakers don't typically stop between words and it would be unbearable if they did!

(4.5) The word is errr probably the ummmmmm most familiar linguistic concept to … non-linguists.

When hesitations occur within a word, they are usually marked by repetition as well as place-holding sounds, akin to *err* and *umm* as in *ling errr … linguistic*). This potential to pause shows native speakers have feelings for what is a word, but it doesn't hold universally, as languages differ in this regard.

In addition, there are the grammatical properties. You might also think about the word as a grammatical unit and this has three aspects to it. One has to do with mobility. It is not possible to rearrange the order of the segments that make up the word, but it is usually possible to move the word itself around in the sentence. *Probably* is transportable, but by itself *-ly* is not.

(4.6) Probably, the word is the most familiar linguistic concept to non-linguists.

The word, probably, is the most familiar linguistic concept to non-linguists.

The word is the most familiar linguistic concept to non-linguists, probably.

Many languages have fairly flexible word order, though English isn't one of them; hence some of these versions might come across as a little awkward. Not all English words have the same mobility as *probably* either, and the little (more abstract) words such as *the* are not free to move by themselves at all. Also problematic is the fact that strings of words can behave as single units with respect to movement. Take our earlier example: *the most familiar linguistic concept to non-linguists*. For a change of emphasis, it can be transported to the start of the sentence:

(4.7) The word is probably the most familiar linguistic concept to non-linguists.

The most familiar linguistic concept to non-linguists is probably the word.

The movement of strings of words is something to which we return in Chapters 9 and 13.

Clearly, mobility alone is not sufficient to establish wordhood, but there is another aspect to the internal structure of words: non-interruptibility (if interruptibility is itself a word!). We have already seen that words have cohesion: speakers don't normally pause in the middle of them and they don't insert material inside them either (ignoring for the moment curiosities like *fanbloodytastic*!). So, the expressions that can be used to modify the word *non-linguists* can only be inserted at the edges of this word:

(4.8) those knowledgeable and well-read non-linguists in this country

Finally, words can stand alone. To our question, *What is the name of the scientific study of language?* you would reply *Linguistics*. You might even create a situation in which a speaker might say, *Did you say a or the concept?* And the answer could be *the*.

To conclude, there are many clues that help us to identify words and these are more or less sufficient for identifying well-behaved words in a language like English. As we've already seen, though, not even all English words fall in line: *the* is clearly a very different linguistic beast from examples such as *familiar* and *concept*. We need to identify different types of words and we also need to recognise different types of meaningful units that are smaller than words. These are topics for Chapter 5 and we only touch upon them here.

Defining words – polysynthetic languages

Some languages are notorious for their complexity at the level of the word, in particular the **polysynthetic** languages (from Greek *poly* ‘many’ and *synthesis* ‘placing together’). Have a look at the following two examples from **Chukchi** (spoken in Siberia) and **Yupik** (spoken in Alaska). They have been broken down into their meaningful parts in (b) below. Don’t worry about the labels for each of these parts; just marvel at the complexity of these ‘words’.

(Chukchi)

(a) *təmeyŋəlevtəpəγtərkən*

‘I have a fierce headache’

(b) *tə-meyŋə-levtə-pəγt-ərkən*

I-great-head-pain-imperfect aspect

(Yupik)

(a) *angyaghllangyugtuq*

‘He wants to acquire a big boat’

(b) *angya-ghlla-ng-yug-tuq*

boat-augmentative-acquire-desiderative-he

Where English has sentences, these languages often have single words. Peter Stephen Du Ponceau (1819), who invented the word *polysynthesis* in the early 1800s, described the word-building characteristics of such languages: ‘the greatest number of ideas are comprised in the least number of words’. Obviously, words in these languages are very different from words in English or, indeed, any of our other case-study languages. So, what makes us even want to call such examples as these words? The question of determining word boundaries is problematic enough in languages such as English; in these polysynthetic languages, it is a very hot topic.

4.2 Lexical versus grammatical words

A man never feels more important than when he receives a telegram containing more than ten words.

George Ade (American humorist) (1905)

The internet has changed the price of communicating with words, but in the years when tele-graphy was the quickest and cheapest way to send messages over long distances, words were not cheap; every word used added to the price. In one study, more than a third of the telegrams sent were under ten words long and less than a fifth had more than fifteen words. A common strategy was to reduce messages to their bare essentials. Following is a pair of examples from a telegram writing instruction manual – which one would you prefer to pay for? How are the words saved?

(4.9a) ‘The enemy has not yet been met or even seen on account of the entanglements thrown up during the night.’

(4.9b) ‘Enemy unmet unseen account entanglements upthrown night.’

Composing telegrams, and to a lesser extent text messages and tweets, pushes writers to focus on words carrying the main semantic content (the **lexical words**), and leave out non-essential words (particularly the **grammatical words**). Consider the first sentence of this chapter:

(4.10) The word is probably the most familiar linguistic concept to non-linguists.

Lexical words such as *word, probably, familiar, linguistic, concept* and *non-linguists* are referable; that is, they relate to aspects of human experience that exist in the real world, although they differ in concreteness – you can touch a non-linguist, while the others refer to concepts rather than physical things. The membership of this class is open-ended, and it involves word classes such as nouns, verbs, adjectives and adverbs (we look at these more carefully in Chapter 5). Languages are acquiring new lexical words all the time and we examine how they do this below.

In example 4.10, the grammatical words do not relate to items in our experience, but rather deal with the relationships between lexical words; that is, they tell us something about them. The little word *the* merely tells us that the expressions *word* and *concept* are defined in some way – in this case, they are known to us and form part of our shared general knowledge. The verb *is* is semantically empty, but its form tells us that we are dealing with the present tense and that the item involved is singular. The word *is* has a linking function to connect the subject *the word* with the rest of the sentence and, in this way, fulfils the requirement of English grammar that every sentence has a verb. *Most* qualifies the adjective *familiar*. Core adjectives are gradable, and this can be indicated by the grammatical words *more* and *most* (*more/most familiar*). Finally, the grammatical word *to* indicates the relationship between *the most familiar linguistic concept* and *non-linguists* (marking *non-linguists* as a kind of recipient).

Grammatical words are far fewer in number than lexical words and they can be exhaustively listed in any grammar; they include prepositions, pronouns, conjunctions and articles (the mysteries of which will be revealed in Chapter 5). In abbreviated styles such as telegrams and text messages, they are the words that are usually either abbreviated in some way or omitted entirely:

(4.11)

arriving 6.30

b there 7 mins

cn u send parcel

4.3 Formulaic expressions (or lexical chunks)

> Words belong with other words not as an afterthought but at the most fundamental level.
>
> Alison Wray (2002), *Formulaic Language and the Lexicon*

In Chapter 1, we described how speakers of a language have an infinite capacity to produce and understand novel utterances. Recall Stephen Fry's memorable sentence, 'Hold the newsreader's nose squarely, waiter, or friendly milk will countermand my trousers.' At the time he spoke this sentence, he was fairly confident that nobody had ever produced this utterance in the history of human communication. While the creative potential of language users is undisputed, when we examine what speakers actually produce on a day-to-day basis, the results are not as fresh as you might imagine. A lot of our utterances involve prefabricated chunks of language – sequences of words that have been stored and retrieved as single units from our memory banks.

Linguists have identified many different types of ready-made language. One type involves short phrases that function like individual words because their parts can't be separated or varied. They include tight packages such as *by and large, so far so good* and *once and for all*. Many of these are idiomatic, which means the meaning can't be predicted from the sum of the parts and the parts cannot be altered in any way. Take the **idiom** *bite the dust*. As soon as you reword this to produce, say, *the dust was bitten*, the literal meaning surfaces and the idiomatic meaning 'to die' disappears.

Other types of fixed phrases are institutionalised expressions. These are short sentences that are also fixed and immutable (*long time no see, all hell broke loose, not happy Jan*). This group includes proverbs, quotable quotes, aphorisms, catchphrases, and the strings of words that we trot out for special events such as greetings and leave-takings – they also include those politeness routines (the so-called magic words) that children are required to learn ('Don't forget to say *Thank you for having me*').

Then there are those linguistic routines that allow some variation but are restricted in what they allow. Take the phrase *as far as I* – this can be completed: *as far as I know/can tell/can make out/am concerned* and so on. You could also include here those words that are optionally, but frequently, associated, such as *blissfully ignorant, fresh-faced youths* and *astronomically expensive*. There are also fixed sequences of words that function like support structures for a larger message; these come in various forms, but they basically provide a kind of framework for building sentences. Examples are *I'm a great believer in ——*, *Let me begin by ——* and *My point is ——* .

These examples are all publicly recognised ready-made phrases, but expressions can also be tailor-made for individuals, and involve speakers' own personal linguistic routines and patterns. Each of us has our own set of fixed expressions.

Slips of the tongue often involve these kinds of fixed phrases; that is, speakers mean to say something but get derailed in the process and end up coming out with something that is a more familiar pattern. Word-association tests (the 'say the first word that comes into your head' type experiments) also show that people are sensitive to word collocations. Responses often involve set phrases such as *sheer coincidence* or *blind date: sheer* triggers *coincidence* and *blind* triggers *date*. We explore these associations further in Chapter 16.

Collocation research and learning Japanese

One of the challenges of learning a new language is gaining a sense of which words belong together. Using them may be automatic for native speakers, but these turns of speech pose real problems for learners and mark them out as non-native speakers; they can have a significant impact on their success in communication with others. Imagine a device that could help you write in a second language by spotting unconventional collocations. Researchers in Japan are designing a tool that checks for mismatches between nouns and verbs by comparing learner-generated texts with a reference corpus to see whether native speakers use the string. To do this, the system takes each word in an unconventional noun–verb pairing (for example, *yume wo suru* ‘to do a dream’) and identifies a range of alternative forms more frequently found for each word. The corpus used by researchers for this tool is a set of data from a social network platform that learners use to write diary entries, which are corrected by native speakers of Japanese. Typical corrections of the verb *suru* ‘do’ in this corpus include *ukeru* ‘accept’, *hajimeru* ‘begin’, *miru* ‘see’, *kaku* ‘write’ and *iu* ‘say’. This includes the correct collocate for ‘dream’: *yume wo miru* ‘to see a dream’. In testing, researchers found that the system was particularly valued by beginner and intermediate learners, who reported that the tool was especially useful when learning new words.

The frequency of these different fixed expressions in language is quite remarkable. One researcher working on a UK corpus of spoken language calculated that 70 per cent of the running words formed a part of recurrent word combinations of some kind. Studies of other varieties of English have yielded similar findings. Much of our everyday communication isn’t terribly new after all!

4.4 Magical words: Names

> People always grow up like their names. It took me thirty years to work off the effects of being called Eric. If I wanted a girl to grow up beautiful I'd call her Elizabeth, and if I wanted her to be a good cook I'd choose something like Mary or Jane.
>
> George Orwell (originally Eric Blair), letter to Rayner Heppenstall, 16 April 1940

Throughout history, people have attributed supernatural powers to names. In ancient Egyptian mythology, Isis gained power over the sun god, Ra, because she persuaded him to divulge his name. Many of you will be familiar with the European folk tales about the evil character variously called Rumpelstiltskin (Germany, parts of England), Terry Top (Cornwall), Tom Tit Tot (Suffolk), Trit-a-Trot (Ireland), Whuppity Stoorie (Scotland) and Ricdin-Ricdon (France). The story is always the same – the discovery of the villain's name destroys his power. In some societies, it is acceptable to know a personal name provided the name is never spoken. In others, the name can never be uttered by its bearer, but is freely used by others.

As an interesting illustration of this power, personal names have been or still are taboo among peoples in many parts of the world. Sorcerers can do harm to a person if they are in possession of that person's true name. A name is regarded as a proper part of a name bearer, not just a symbol but the verbal expression of an individual's personality. For this reason, in many languages, a name is an **inalienable possession**; like body parts (and other properties of personal representation, such as mind, spirit, soul, shadow and products of bodily functions), names don't exist apart from the possessor (so are different from, say, a person's shoes or house), and this can have repercussions for the grammar. For instance, in the Aboriginal language Pitjantjatjara, a name is treated like a body part and unlike a boomerang: contrast (a) and (b) with (c) in example 4.12.

(4.12)

(a) *ini ngayu-nya*
name I-accusative
'my name'

(b) *mara ngayu-nya*
hand I-accusative
'my hand'

(c) *kali ngayu-ku*
boomerang I-possessive
'my boomerang'

You can see here that speakers use the ending *-ku* to mark the possessor of something like a boomerang (as in c), but not to mark the possessor of a body part or a name (as in a and b). Our case-study language Gurindji also treats entities such as body parts and shadows as inalienable nouns (i.e. no marking); however, the Gurindji word for *yini* ‘name’ is an alienable noun (no different from, say, boomerang or car).

In Mali, names are also treated in the same way as body parts and close kin. These kinds of words must always carry a possessive pronoun, even when they are not used to refer to the name/body part/close kin of anyone in particular. Compare (a), (b) and (c) with (d) in example 4.13.

(4.13)

(a) *angēt-thēp-ki*
their = name- feminine.singular
‘(someone’s) name’
(b) *angēt = keng-ki*
their = tooth- feminine.singular
‘(someone’s) tooth’
(c) *angēt = nanēk-ki*
their = mother- feminine.singular
‘(someone’s) mother’
(d) *uat-ki*
basket-feminine.singular
‘basket’

Because true names are so closely associated with their name-bearer as to be a proper part of that person, in some societies there are strict taboos preventing two living persons from going by the same name. Furthermore, true names are often secret, so that substitution names are necessary for public naming and addressing. In many places, names of the dead are (or until recently were) taboo. Sometimes the ban extends to those personal names that the dead person may have given to others. Violations of such taboos are believed to cause misfortune, sickness and death; they may also cause offence to living descendants.

In many traditional Australian Aboriginal communities, any kind of vocabulary item, including grammatical words, can be outlawed if it is the same as, or phonetically similar to, the name of a recently deceased person. Replacement vocabulary is created using other expressions from the language's own repertoire; this process might involve meaning shifts of existing words or circumlocution (long-windedness). Borrowing from a neighbouring language and, in some cases, using a hand sign or gesture are other strategies that are used widely. Some languages have special vocabulary items to be used in place of proscribed words; sometimes a kind of 'whatsitsname' at the end of a word or 'whatchamacallut' at the start of a word is reserved especially for the purpose of name avoidance. For example, in the case of Gurindji, if someone dies, their name and any nicknames are taboo, and any similar-sounding word is also taboo. The person is then either referred to by kinship terms – for example, 'Granny' – or their skin name – for example, *Nangala*, often with a deceased ending, such as *Nangala-yayi*. They may also be referred to by a 'no name' term, *yini-murlung* (name-not) or *call em-murlung*, or even sometimes just *call-em* ('call-them').

There are probably few Westerners today who would admit to still believing in the supernatural powers of names, yet the practice of onomancy (interpreting names as omens) is alive and well. Think about the occasional punter at the track who selects a horse simply because its name has some special significance – perhaps it links with a recent event or significant person. The keywords 'name' and 'magic' typed into any web search will uncover instructions on how to cast a spell on someone using their true name, how to find the luckiest name for a newborn baby, or how to find the perfect soulmate using names and numerology to compute the love, sexual and intellectual compatibility of partners. There are also sites that will assist in finding the meaningful anagrams of names and ways of interpreting these to define one's destiny.

Even those who disparage this kind of superstitious carry-on manifest forms of behaviour that strongly suggest a belief in some sort of magical relationship between a name and the individual who bears it. Many experience difficulty when they have to say their own name out loud; this is not confined to young children. Is it vestigial superstition about revealing one's real name, or is it simply that names are so inextricably bound up with our innermost selves that we feel embarrassed to reveal such an intimate detail to strangers? When someone misspells our name, they touch a sore spot. From a very early age, children learn the effectiveness of name-calling, often upsetting their peers by focusing on physical aspects (*square eyes, fatty*) or playing on the rhyme of their names (saying for no good reason *Good on ya Tonya*, or changing *Burridge* to *Porridge*). We are our names and insulting names do hurt.

A presidential-sounding name

In a 1998 review of successful presidential candidates in the United States, certain phonetic features are common in the names of successful candidates. Two-syllable names are preferred over one-syllable names, as is an initial sequence of consonants containing an /l/ or an /r/, and a /m/, /n/ or /ŋ/ in word final position (sometimes in a cluster with a /p/, /b/, /t/, /d/, /k/ or /g/). Among the dispreferred features are words ending in fricatives (sounds like /f/, /v/, /s/, /z/).

In February 2016, John Oliver referred to the importance of the sounds of names in election campaigns:

> I know his very name is powerful … Trump, it's the sound produced when a mouthy servant is slapped across the face with a wad of thousand-dollar bills. Trump. It's the sound of a cork popping on a couple's champagne anniversary, the day that renovations in the wine cellar were finally completed.

As Oliver notes, in an earlier generation, the Trump family had changed its name from Drumpf – which doesn't have quite the same ring about it. Oliver goes on:

> Drumpf is much less magical. It's the sound produced when a morbidly obese pigeon flies into the window of a foreclosed Old Navy. Drumpf! It's the sound of a bottle of store-brand root beer falling off the shelf in a gas station minimart … So if you are thinking of voting for Donald Trump, the charismatic guy promising to 'Make America Great Again', stop and take a moment to imagine how you would feel if you just met a guy named Donald Drumpf … Would you think he would make a good president or is the spell now somewhat broken?

The significance of naming in modern Western society shows up in other ways. Name-dropping is believed to give social clout. People dispose of their names and adopt new ones to promote a better public image. People can even sell their names, as when someone endorses a commercial product. Those who buy the item expect to pay more because of the name inscribed on the goods. We have common phrases such as *make a name for oneself; have a good name; bring one's* *name into disrepute*; and *clearing one's name*. We will see in a moment how personal names can lose their capital letter and enter the general lexicon as household words; eponymic phrases arise spontaneously in everyday language (for example, when *obama* popped up as a slang word for 'cool'), but they are typically short-lived, as in the case of *obama*.

All this ostentation appears in striking contrast to the behaviour of those in societies where there are strict taboos on naming and where true names remain secret. How different our media-conscious twenty-first century seems: everyone is out there striving to *be a name*. What we see here, though, is simply two different sides of the same coin; whether they are revealed or concealed, names are highly charged vocabulary items: they carry power and mystique.

Language and the law: What makes a name?

For a variety of reasons, many people seek to change their names. In Australia alone during 2010–11, there were more than 50 000 changes (and this figure doesn't include the 900 000 other people who sought to have an additional name). Of course, it has to be determined that the reasons are legitimate (you can't create an alias for the purpose of committing fraud or escaping debt), and you also can't change your name to whatever you wish. This is where linguists occasionally come in – to provide expert opinion on whether something is a viable name, whether the new name might cause confusion in the community (perhaps it resembles an official title or rank) and whether it might be considered to be obscene or offensive. Names that have been rejected in the past include Batman, Jesus Christ, Prime Minister, Pappa Smurf, Motorcycle Feelgood, Smart Arse and Cu L8 r. One of the most celebrated cases in Australia involved someone seeking to change his name to: Prime Minister John Piss the Family Court and Legal Aid.

Linguists are also called on to provide expert opinion when a trademark infringement has been alleged. Brand names as well as personal names can become household words (e.g. *biro* for ballpoint pen), and while some corporations don't mind this, they do mind when other corporations move in on what they see as their trademark territory. When, in 1988, Quality Inns International planned to launch a new chain of economy hotels under the name 'McSleep Inns', the McDonald's Corporation fought against this on the grounds of trademark infringement and unfair competition. Of course, the hotel company denied this, but in the end McDonald's triumphed.

4.5 Lexical addition

> Three quarks for Muster Mark!
> Sure he has not got much of a bark
> And sure any he has it's all beside the mark.
>
> James Joyce (1982 [1939]), *Finnegans Wake*

We now move to a fascinating aspect of the lexicon: word creation. Words are seldom created from scratch and it is hard to find examples of true coinages (words not built on pre-existing elements). Examples usually involve proper names, such as the trade name *Kodak*. It first appeared in 1888 and, according to George Eastman who came up with the name and the product, it is not based on any other word – the inspiration was his love of <k>, which he described as 'a strong, incisive sort of letter'. Eastman's biographer, however, reports that this fondness for <k> probably came from his mother's name, Kilbourn, so the two appearances of this letter in *Kodak* are not entirely accidental. The technical word *quark* is another potential coinage. It first appeared in James Joyce's book *Finnegans Wake* and was later taken over by physicists to describe some sort of elementary particle of matter (not *quark* 'the call of a gull'!). But, who knows what particular associations this word might have had for Joyce?

Whether words are derived on the basis of already existing words, or coined from sounds and sound patterns, the resulting forms must conform to the rules of the language concerned (which is why Trump's *covfefe* would never have been one of the best words). Which of the candidates listed in Table 4.1 could be acceptable new words in English? Can you guess at what they might mean?

Table 4.1 Candidate new words in English

Coinage	**Derivation**
blatch	dis-invite
frelp	un-purchase
splang	re-knit
flip	phon(e)-able
phwomma	phon(e)-ability

As we mentioned, words rarely arise out of the blue, and coinages such as *blatch, splang* and *frip* are actually quite rare. It is much easier for people to learn a new word if it is based on elements they already know. Consider how much harder it is to 'guess' at the meanings of the coinages in Table 4.1 than the derived forms.

An important restriction on new words in any language is that people have to be able to pronounce them (something we explore in more detail in Chapter 8). The possible words *blatch, splang* and *frip* are acceptable because they are consistent with the sound patterns of English. This is clear if we compare these words with the following existing words: *black/catch, splat/sang, fritter/trip*. In comparison, *fnelg* and *phwooma* are not acceptable in English because they do not follow the sound patterns of the language. This is because the consonant clusters suggested by *fn* and *phw* are not available in English (well, apart from curiosities such as the interjections *fnarr!* and *phwoooar!*).

In terms of their meanings, the only requirement really is that new words should be useful in some way. You might think about what areas of life have seen the development of new words in the last twenty years (computers, types of music, lifestyle and so on). In the following discussion, we examine some of the word-formation processes that are behind these creations; we emphasise here that these processes are found in languages around the world (though there are usually differences in significance and liveliness when it comes to word creation). As usual, we will start with English before moving to a cross-linguistic perspective, in this case provided by German.

Acronyms

Acronyms are words formed from the initials of other words. For example, *mooc* comes from 'massive open online course' and *fomo* from 'fear of missing out'. The actual label 'acronym' came into being in the 1940s and is based on the Greek words *acro-* 'tip, or point' and *onym* 'name'. Technically, for something to be an acronym, the resulting word has to be pronounceable like ordinary words in the language. So, examples that are pronounced as strings of letter names, such as BYOD ('bring your own device'), are not acronyms but rather **initialisms** (or **alphabetisms**). Some words manage to fit into both categories. The now international word OK is an abbreviation and an acronym because the letter names also make a pronounceable English word (*okay*).

Occasional examples are based on successive syllables of *single* words. They include a few well-known examples such as *TV* from 'television' and *PJs* from 'pyjamas'. Other variations involve speakers taking more of the words than simply the initial letters. This might mean, for instance, using the first consonant and vowel. Usually this is to make the acronym pronounceable, as in the case of *sonar* from ‘sound navigation and ranging’, which begins with the first two letters of the leading word in the sequence.

Table 4.2 shows some of the everyday examples of German acronyms. German speakers, like their English counterparts, feel they are swimming in a sea of acronyms these days, but in both cases the majority of expressions remain peripheral to the languages. They serve either as proper names or are specific to certain occupations, and are never likely to become a part of the general lexicon.

Table 4.2 Examples of German acronyms

Word	**Source**	**Literal translation**
Bund	*Bund für Umwelt und Naturschutz Deutschland*	‘Association for environment and nature conservation Germany’
MOF	*Mensch ohne Freunde*	‘people without friends’
PIN	*Persönliche Identifikationsnummer*	‘personal identification number’
UFO	*Unbekanntes Flugobjekt*	‘unidentified flying object’

Once they have been around for a while, acronyms enter the language as ordinary words, such as *sonar* and German *Bund* (*Bund* has an initial capital here because all German nouns do). Even where this hasn’t happened, the original source words are usually forgotten – *HIV* stands for ‘Human Immuno-Deficiency Virus’, yet people commonly refer to the ‘HIV virus’ (compare *ATM machine* ‘Automatic Teller Machine machine’ and *PIN number* ‘Personal Identification Number number’). Examples like *PIN-Nummer* in German have given rise to an acronym to describe the practice: *RAS-Syndrome* (*Redundantes-Akronym-Syndrom-Syndrom* ‘redundant acronym syndrome syndrome’), a joke that works equally well in English.

Often, you can have expressions that are more like reverse acronyms. Consider the two words *wimp* and *macho*. Both are technical acronyms in physics – *WIMP* standing for ‘weakly interacting massive particle’ and *MACHO* standing for ‘massive astrophysical compact halo object’. Yes, it could be a happy coincidence that the initials in these phrases formed such apt and cute-sounding words as *wimp* and *macho,* but more likely the creators of these ‘backronyms’ fudged and fiddled until they came up with the right sequences of words.

Shortenings

Shortenings or clipped words are simply abbreviated forms that end up replacing the longer versions (*mob* has replaced *mobile vulgus* ‘movable or fickle common people’ and *bus* has replaced *omnibus* 'motor vehicle for paying passengers'). Often, the longer and shorter forms coexist as stylistic variants (*dis* alongside *disrespect*), and it can happen that the meanings end up diverging (*hussy* versus *housewife, stroppy* versus *obstreperous, grotty* versus *grotesque* and more recently [mobile] *app* versus [computer] *application*). The use of the shorter colloquial forms in casual contexts often leads them down quite different semantic tracks, thus creating two radically different words as the long and the short part company. Special cases of shortenings arise when initial unstressed syllables are lost (*cos/coz* from *because*), and this can also give rise to separate words with different meanings (*fence – defence; cute – acute; squire – esquire; ticket – etiquette*).

As in the case of English, shortening in German is not a terribly important means of adding to the lexicon; mostly the clippings coexist with their longer counterparts. Some examples of German clippings are shown in Table 4.3.

Table 4.3 Examples of German shortenings

Word	**Source**	**Translation**
Ober	*Oberkellner*	'head waiter'
Dia	*Diapositiv*	'slide, transparency'
Rad	*Fahrrad*	'bicycle'
Uni	*Universität*	'university'

Bro and *brother* go their separate ways

Bro has been a written abbreviation for *brother* since the 1500s, and in the plural form *Bros*, it has been part of company names for some time (*Warner Bros*). But, it was in the early 1900s that the abbreviation really took off and started to break away from its older sibling *brother*. Originally African American usage, *bro* appeared as a familiar form of address, particularly among close-knit groups of male peers (*What's up, bro*?), and then as a more general term for 'guy' (much like *dude*). *Brodom* is now identified as a subculture of young men (like *brotherhood*, the *Urban Dictionary* explains, but 'way stronger'), and as part of this subculture we are seeing a flurry of *bro*-based neologisms (dubbed *brocabulary*). The life cycle of slang is always a puzzle. If they fill a need, expressions can stick around – *bromance* 'intimate and affectionate friendship between men' has been in dictionaries for a while now. But it is the nature of slang that there is always a turnover. Voguish *curlbro* (used of gym-goers who focus on training their arms) and *brogrammer* (the stereotypical male computer programmer) enjoyed some popularity, but this was short-lived.

Compounding

Compounding is a word-formation process that is found extensively in languages around the world; indeed, word-formation expert Laurie Bauer claims there is no known language that does not have compounds. The process involves the combination of two (occasionally more) free-standing forms. What is crucial is that the new word has a meaning that is in some way different from that of its elements (it is usually more specific); compare *black boards* with *blackboards* (which can be green). Compounding has always been an important source of new lexical items for Germanic languages, such as English and German, and continues to be. New creations include *toxic debt, fanta pants, moral hazard, helicopter parenting, lifestreaming, guerrilla gardener, water footprint, mouth grills, Wii shoulder* and *seagull manager* (‘a manager who flies in, makes a big mess and lot of noise, and then leaves’).

The ‘one-wordedness’ of compounds becomes obvious when we look at the way they behave in the grammar. For example, the compound *still-life* has a regular plural form (*still-lifes*) compared to the irregular plural of *life* (*lives*). Similarly, you would say *jack-in-the-boxes* rather than *jacks-in-the-box*, which means you are thinking of *jack-in-the-box* as a fused word. You can also tell by the stress pattern. There is a heavy stress on the first syllable and the following parts are all squashed together: ***Jack***-*in-the-box*. Meaning is also significant. Take, for example, the wide range of relationships possible in noun–noun compounds: in *flu symptoms*, flu causes the symptoms, so A causes B. However, in *flu virus* A is caused by B, and in *flu shot* A is prevented by B. In *orange juice* B is made of A; in *rubbish bin* B contains A; in *frogman* B resembles A; in *needlework* B is made with A, and so on. Particularly interesting are those compounds with highly idiosyncratic meanings, such as *sex appeal, bag lady, spaghetti western* and *black sheep*. The fact that we are thinking of these expressions as single units is obvious from their unified meanings. Similar examples may be found in German, as we show in Table 4.4.

Table 4.4 Examples of German compounds

Word	**Parts of compound**	**Literal translation**
Handfläche	*Hand* ‘hand’ + *Fläche* ‘surface’	‘palm of the hand’
Kopfkino	*Kopf* ‘head’ + *Kino* ‘cinema’	‘imagination’
Spiegelei	*Spiegel* ‘mirror’ + *Ei* ‘egg’	‘fried egg’

Sprachgefühl	‘language intuition’	*Sprache* ‘language’ + *Gefühl* ‘feeling’
Verbesserungsvorschlag	‘suggestion for improvement’	*Verbesserung* ‘improvement’ + s + *Vorschlag* ‘suggestion’
Schweinefleisch	‘pork’	*Schwein* ‘pig’ + e + *Fleisch* ‘meat’

These examples show that German compounds sometimes have a connecting element between the parts (here *-s-* and *-e-*). They also show that, in the German writing system, parts of the compound are joined to form one long word, and sometimes very long words. One of the authors of this book is the proud owner of a German kitchen device known as the *Eierschalensollbruchstellenverursacher*, referred to colloquially as the ‘Clack’ (*Eier-schalen-soll-bruch-stellen-verursacher* ‘egg-shell-necessity-break-spot-causer’).

In varieties of English (such as bureaucratese), you occasionally encounter similarly complex clumps of words, although they are not usually joined (*prototype crisis shelter development plans; occupational choice–vocational interest congruency*). But conventions for writing even straightforward English compounds vary. Sometimes they are written solid (*brocurl*), and sometimes separated (*bro hug*) using spaces or hyphens (*bro-husband*). Dictionaries don’t agree on the conventions they adopt, although on the whole American editors abandon hyphenation more quickly than others do. Often, it depends on the length of time that a compound has been around. Well-aged compounds will typically lose their hyphens and appear as one word, but speakers often disagree. The iced dessert has been on our menus since the 1600s, but we still can’t agree on how to write it – is it *icecream, ice-cream* or *ice cream*?

Familiarity not only removes hyphenation in well-established compounds; it can bring about spectacular reduction as the elements of the compound become more and more closely associated. Look at what has happened to the pronunciation of *cupboard*. Here the spelling has preserved the original compound, but this is not always the case – only the real word enthusiast will be aware that *nostril* began life as a compound (Old English *nosþyrel* ‘nose thirl [=hole]’). More unusual compounds don’t show the same sort of reduction. Compare the full pronunciation of an infrequent word like *handstroke* with the squished sounds in the more common *handkerchief*.

Reduction in Auslan compounds

As linguists working on casual speech processes have pointed out, when movements of the tongue, lips and mouth (gestures of articulation) become automatic, this triggers a reduction in the timing and magnitude of these gestures. The speed of execution automatically increases, the transition between sounds becomes more fluent and movements might even overlap. Sign language users are no different. As Johnston and Schembri describe, the various formational changes affecting Auslan compounds are generally reductive. There is a smooth and fluid movement between the two elements, and the duration of the compound is shorter than that of two individual signs. The movement in the first sign is condensed – for example, when it is an individual word, LITTLE is made with repeated or twisting contact, but when it is part of the compound word for CHILD (LITTLEˆSHORT), it is produced with a single contact. Furthermore, where signs involve repeated movements, as in the compound word for PARENTS (MOTHERˆFATHER), they are contracted to a single movement.

When in familiar territory, we take short-cuts. Some might see this reduction as laziness, but equally you could view it as efficiency. Signers and speakers are saving on time and effort for those words they know well, and they are conserving their co-communicators’ time and effort, too. After all, familiar words are easier to access than less familiar words, so they don’t need the same amount of detail.

Affixation

Affixation is an important way of creating new words in both English and German. It is a similar process to compounding except it involves parts of words that can’t stand alone (we will have more to say on these in Chapter 5). Affixation includes **prefixes**, which are added at the beginning (*un-, mis-* and *non-*, as in *unhappy, misfortune* and *non-linguist*), and **suffixes**, which are added at the end (*-ish, -ness* and *-ic*, as in *blackish, happiness* and *linguistic*).

Some affixes are highly **productive** and achieve almost voguish popularity; since the 1950s, there has been an explosion of *-ise* words (*prioritise, personalise, diarise* and so on). English has many different prefixes and suffixes, and new ones are always appearing on the scene. McDonald’s has provided the new *Mc-* prefix, giving us a spate of McCoinages such as *McJobs* ‘unstimulating, low-paid jobs with few prospects’, *McBesity* ‘fatness relating to the consumption of junk food’, *McMansions* ‘housing developments made up of large, cookie-cut houses’, *Baldy McMullet* ‘mullet worn with very little or no hair on the front and sides’, *McGarbage* ‘waste created by disposable dishes, cutlery and packaging', *McBooks* 'quickly produced mass-market books', *McMovies, McShopping, McMedicine, McPrisons* – even *McGod* 'The god of TV evangelists'. Not all of these will survive (some have already disappeared), but the flourishing of such creations shows the success of this prefix. Clearly McLanguage has moved beyond products explicitly identified with the fast-food giant (*McMuffins, McNuggets, McCafé*).

Just as in English, German draws on many prefixes and suffixes to expand its lexicon. The languages even share some of these affixes; for example, *be-* in *befriend* and *besprechen* 'discuss', or *-er* in *singer* or *Sänger* 'singer'. Table 4.5 shows a few more examples.

Table 4.5 Examples of German affixation

[illegible] **Translation**
[illegible] 'black'
[illegible] hol'
[illegible] 'secret'
[illegible] g'
[illegible] eak'

How long can English words get?

If we put aside chemical compounds and nasty lung diseases (such as *pneumonoultramicroscopicsilicovolcanokoniosis* 'a lung disease caused by inhaling very fine dust'), *floccinaucinihilipilification* at 29 letters is the current champion (defined as 'the action of estimating something as worthless'). As linguist Steven Pinker points out, though, we can create *floccinaucinihilipilificational* 'pertaining to the estimation of something as worthless'. That has two more letters. Then again, *floccinaucinihilipilificationalise* 'to cause something to pertain to the estimation of something as worthless' has another three. Or how about *floccinaucinihilipilificationalisation* 'the act of causing something to pertain to the estimation of something as worthless'? But wait, from that we can build *floccinaucinihilipilificationalisational* 'pertaining to the act of causing something to pertain to the estimation of something as worthless'. And so we can go on creating longer and longer words – there is no theoretical limit.

Affixation, of course, has a practical threshold, to do with the limited brain space available to keep track of impossibly long expressions. But, theoretically, words and parts of words can combine and recombine in this way to form an infinite number of different words. German and English illustrate the structural complexity and creativity that distinguishes our communicative behaviour from that of animals. The infinite capacity to express and understand meaning is not found in the language of any other species. You have already encountered examples of the remarkable inventiveness of our language in Chapter 1, and you will encounter different aspects again and again in this book.

Backformation

Backformation is the opposite strategy to affixation. It can happen that words exist with prefixes and suffixes, but not without them. The forms *inept* (from the Latin *ineptus* ‘unsuitable, tasteless’) and *dishevelled* (from the Old French *deschevele* ‘stripped of hair’) were borrowed wholesale into the language, and their negative prefixes *in-* or *dis-* were part of the package. Neither had a simple positive form in English. Some time during the middle of the nineteenth century, speakers took *inept* and created the positive counterpart *ept* ‘competent, adroit’. Occasionally you can hear speakers use *shevelled* to mean ‘neat, tidy’. This form hasn’t yet made it into the dictionaries, but one term that has gained acceptance is the verb *dishevel*. There is the adjective *dishevelled*, so surely there is the related verb *to dishevel*. We use back-formation to fill the vacancy. This process has given us many verbs, including *afflict, enthuse, laze, aggress, grovel, televise, manhandle, eavesdrop, househunt* and many more.

Although backformation is also a minor word-formation process in German, examples like those in Table 4.6 are easy to find.

Table 4.6 Examples of German nouns backformed from adjectives

Word	Translation	Source	Translation
Neugier	‘curiosity’	*neugierig*	‘curious’
Sanftmut	‘gentleness’	*sanftmütig*	‘gentle’
[illegible]	‘grumpiness’	*mürrisch*	‘grumpy’
Tiefsinn	‘profundity’	*tiefsinnig*	‘profound’

In some cases, speakers remove something that they believe to be an affix (but it isn’t) and create a new word. In the case of the verb *to verse* (as in *New Zealand is versing Australia*) speakers have reanalysed the word *versus* as *verses*. So, they have removed what looks to be a verb ending and created a new verb. Such backformations are more likely to occur with very strongly entrenched patterns, and they also have the effect of filling an apparent void – if there is a noun (*butler*) then there should be a verb (*to butle*); in fact, there once was, but it didn't survive (perhaps because butlers didn't). In the same way, the word *beg* was created from *beggar*, the final *-ar* wrongly interpreted as the same *-er* suffix as in *bake* and *baker* (the word *beggar* comes from Beghard, a member of a medieval Christian brotherhood).

Blends

Blends are new words that have been created from the combination of two (and occasionally more) existing words. The new portmanteau word then incorporates meaningful characteristics from both – for example, *fantabulous* is a combination of *fantastic* and *fabulous*. Blends don't always involve squishes of this kind, however. They can also combine full words with parts of others. *Shoefiti* 'the practice of decorating over-head wires with shoes' is a blend of *shoe* and *graffiti*. Blending has never been considered a significant method for creating new words in English, but this seems to be changing and linguistics books probably need to revise their descriptions of this word-formation process. Dozens of new portmanteau words are appearing all the time. Furthermore, the nature of the mixing process is changing too. Rather than combining splinters of words, most of the new blends appearing now are more like *shoefiti* – complete words merged with parts of others. It might be that the front is intact, as in *shoefiti*, but sometimes the blend starts with a fragment and is then followed with a full word. American economic and cultural expansion is now often described as *cocacolonisation*. This is a clever fusion of part of *Coca-Cola*® and *colonisation*. It can also happen that the full word appears in the middle of the blend, as in *ambisextrous* (where *sex* is squashed inside *ambidextrous*), and occasionally the overlap is total (*sexploitation* retains both *sex* and *exploitation*). This is also true of *bagonise* 'to wait anxiously at the baggage carousel for your luggage to arrive'. By integrating *bag* and *agonise* in this way both words remain intact (compare earlier *bromance*). All this can be used to great effect in advertising and in product naming (*hangry*, a blend of *hungry* and *angry*, makes regular media appearances).

Blending is not nearly as common in German, but some examples are shown in Table 4.7.

Table 4.7 Examples of German blends

Word	**Source**	**Translation**
Bistrothek	*Bistro* + *Diskothek*	'restaurant with music'
Teuro	*teuer* 'expensive' + *Euro*	slang for Euro
Müllionärin	*Müll* 'trash' + *Millionärin* 'female millionaire'	'millionairess in trash scandal'

Conversion

Conversion changes one part of speech to another without anything being added. In English, conversion usually involves the major word classes: nouns, verbs, adjectives and adverbs. The most common conversions are as follows.

- verb → noun (a guess, a call, a think, a read, a big ask)
- noun → verb (to bottle, to bridge, to trash, to network, to leaflet)
- adjective → verb (to better, to empty, to open, to total [a car])
- adjective → noun (a roast, a weekly, a regular, a given, a nasty).

There don’t appear to be any restrictions on conversion – words can move between classes with extraordinary ease. There are even examples of conversion from minor classes, such as prepositions and conjunctions. *Up* and *down* can be more than just simple prepositions – see, for example, these phrases: *to up the stakes/to down tools* (verb), *on the up/to have a down on someone* (noun), *he’s up today/a down experience* (adjective) and *to go up* and *down* (adverb).

German conversion is far more restricted on account of the fact that German words have more in the way of grammatical endings. Hence affixation is more common (when a noun becomes a verb, for example, it will take on the necessary verbal endings, as in *Scherz* ‘joke’ > *scherzen* ‘to joke, *Salbe* ‘salve’ > *salben* ‘to anoint’). The most common conversions are nouns created from basic verbs, for example, the verbs *kochen* ‘to cook’ and *verabschieden* ‘to say farewell’ may also function as nouns: *Kochen* ‘cooking’ and *Verabschieden* ‘saying farewell’.

Commonisation

Commonisation describes the process whereby proper names extend from specific cases and end up referring generally to the whole class of items. Brand names provide a common source for new words. The word *daks* now refers to a basic pair of trousers or pants (as in *tracky daks*) but originally was the exclusive label of Simpson’s of Piccadilly. The less glamorous *stubbies* were a brand of shorts, but now apply widely to any shorts of that style. It is also common for personal names to enter the language as ordinary words, to extend from the specific case of the celebrated name-bearer and refer generally to the class of phenomena. In protest, we might *boycott* something by refusing to do business (after Charles Cunningham Boycott, a nineteenth-century landlord who was evicted by his own tenants because he wouldn't lower the rents). Ned Lud, who destroyed two knitting frames in the late 1700s, gave us *luddite*, the forerunner of the modern-day *technophobe*. *Sandwiches* are associated with the Fourth Earl of Sandwich, who enjoyed this neat way of eating meat and bread either at the gambling table or at his desk (it's not clear which). We *macadamise* roads (a method devised by John Loudon *McAdam*), we wear *wellingtons* (after the first Duke of Wellington) and test the *voltage* of our batteries (after the Italian physician and scientist Alessandro Volta). Over the years, English has amassed a surprising 35 000 **eponyms** (words derived from proper names). Several examples from German are shown in Table 4.8.

Table 4.8 Examples of German eponyms

Word	**Source**	**Translation**
Röntgen	Wilhelm Conrad Röntgen	'x-ray'
Sadismus	Marquis de Sade	'sadism'
Birchermüsli	Maximilian Bircher-Benner	'muesli'
Knigge	Adolf Freiherr von Knigge	'etiquette manual'

Google.com strikes back

The success of Google is demonstrated by the appearance of the new verb *to google*. To say something like 'I googled it' means to use the Google website to search for information on a certain topic. But, the verb is also well on the way to generalising its meaning. Many now use *to google* to refer to any search, not simply a Google search. This new verb illustrates two word-formation processes we have just discussed, namely, conversion and commonisation; that is, the proper noun *Google* has been turned into a verb, *to google*. In 2002, the verb was judged the 'most useful' new word of the year by the American Dialect Society, and it was officially added to the *Oxford English Dictionary* in June 2006, and to the 11th edition of the *Merriam-Webster Collegiate Dictionary* in July 2006.

However, rather than be flattered by the fact that it has been 'verbed' by this new creation, Google.com has gone to considerable lengths to try to remove the word from dictionaries.

When *to google* loses its capital letter and generalises to mean any search of the internet, the name becomes a household word, as has happened in the case of *hoover, biro, esky, escalator, gunk, xerox* and *kleenex* (and its plural *kleenices*). While we might see this as the ultimate sign of success, the company's lawyers argue that the name loses its value. They would like to stamp out the new verb and have applied pressure to dictionary editors to delete the definition or revise it to take account of the 'trade mark status of Google'.

Reduplication

Reduplication is a repetition process whereby all or part of the stem of a word is reiterated (we'll have more to say on word stems in Chapter 6). The resulting form is a kind of compound. This has been considered a peripheral process in English compared with other languages; nonetheless, over the years it has produced some 2000 words and new ones are occasionally being created (*cray cray* 'crazy' is a recent addition). In English, there are basically three types of reduplicated compounds. One involves repetition of the whole stem (*goody-goody, fifty-fifty, hush-hush*). Many of these are confined to nursery language (*woof-woof, yum-yum, bow-wow*). A second type involves the repetition of the rhyme. Sometimes both stems are existing words (*brain-drain, tin-grin, culture-vulture*), but usually only one, or sometimes none, of the elements is independently meaningful (*argy-bargy, arty-farty, lovey-dovey*) – it can happen that the words become obscure through sound change, or else the words might simply drop out of use (*willy-nilly* derives from the expression *will I nill I*, which is based on the earlier verbs *willen* 'want' and *nillen* 'not want'.) A third (now rather rare) type of reduplicative compound involves some sort of modification of the stem vowel (*shilly-shally, dilly-dally, riff-raff, flip-flop*).

Like English, reduplication in German is not a lively word-formation process. It is largely confined to child language (*wauwau* 'woof-woof '/'dog'), expressive (onomatopoeic) creations (*ticktack* 'tick-tock') and colloquialisms (*ratzfatz* 'very fast'; *muckefuck* 'coffee substitute' and now 'stuff and nonsense'). A few additional instances are shown in example 4.14.

(4.14)

zickzack ‘zigzag’
ruckizucki ‘quickly’
rappzapp ‘quickly’
Mischmasch ‘hotch potch’

‘You mean he’s GONE gone?’ – linguistic cloning

We copy words for all sorts of reasons. For instance, there are those expressions we associate with childhood, as in the examples just mentioned (like, *choo-choo* or *gee-gee*). We might also repeat things to beef them up, as in *I’m OK I’m OK*. This can be for emphasis, to convey a sense of conviction or urgency. Other examples are a little more interesting. Take the following (overheard) example: ‘You mean he’s GONE gone?’ (Note, there is a heavy stress on the first *gone* that contrasts with the second mention.) This person was asking whether the chap had actually gone for good, as opposed to just ducking out for a short while. This is something that linguists Jila Ghomeshi, Ray Jackendoff, Nicole Rosen and Kevin Russell have dubbed ‘contrastive focus reduplication’. As they point out, even entire expressions can be copied in this way. You could imagine someone saying, ‘Did you TALK-ABOUT-IT talk-about-it?’ – in other words, ‘Did you actually talk about it, or just drop it into the conversation?’

In these examples, the copied expression points to the ‘real’ or true meaning of the term referred to. You can usually rephrase the whole thing using modifiers like *real* or *really*. Instead of *he’s GONE gone,* you could say *he’s really gone.* Usually the repetition indicates that the literal meaning is intended. These linguists provide a nice example from a conversation involving a married couple. They had recently separated and were no longer living together. One said, ‘Maybe you’d like to come in and have some coffee?’ When the other partner agreed to this, the first went on to explain, ‘Just COFFEE-coffee, no double meanings’. Of course, *having coffee* is one of the thousands of verbal disguises English speakers resort to when they need to talk about having sex. The repetition here signals that the speaker intends the literal meaning of *coming in for a coffee*, not the euphemistic one.

Borrowing

> We don't just borrow words; on occasion, English has pursued other languages down alleyways to beat them unconscious and rifle their pockets for new vocabulary.
>
> James Nicoll (1990), Usenet group rec.arts.sf-lovers

So far, we have only considered how languages enrich their vocabulary by drawing on resources that are in-house (in other words, processes internal to the language), but clearly external sources also provide an important method of lexical expansion. **Borrowing** is the process whereby a language (or dialect) takes and incorporates some linguistic element from another language (or dialect). Words are the most borrowable part of the language, but all parts of the grammar can be 'on loan' (and we address the issue of language contact in Chapter 12).

As British linguist David Crystal observes, English is 'a kind of vacuum cleaner of language – it sucks in vocabulary from any language it can get'. Indeed, more than 120 different languages have contributed to the lexical coffers of English. The German language makes up only a tiny proportion of this lexical loot, yet there are some significant items among them: everyday words, such as *delicatessen, kindergarten, rucksack, poodle, gimmick, waltz* and *dunk*; food and drink terms, such as *noodle, pretzel, lager, pumpernickel, sauerkraut* and *schnitzel*; and, of course, *frankfurter* and *hamburger* (adjectives formed from the place names Frankfurt and Hamburg). Most of these lexical aliens have been naturalised – they fit the English sound and spelling system (no capital letter on the nouns) and they have been integrated into the grammatical system (for example, nouns take the English plural ending *-s*). Sometimes this assimilation process changes the loan words beyond recognition: *Kris Kringle* (known as *Secret Santa* in some parts of the world) has come a long way from the dialectal German *Christkindl* 'Christ child'. *Hamburger* offers another spectacular illustration. The first part of the word happens to correspond to a type of meat in English, and the word has been 'misinterpreted' as a compound, *ham* + *burger* (even though there is no ham involved). This false analysis has spawned many new compounds, such as *Aussie lamburger, eggburger, cheeseburger, chickenburger, steakburger, vegeburger* and even *burger bits* (for the dog).

Other loans retain more of their Germanness – they have been adopted but not adapted, or at least not yet. For example, German has a handy word to describe excess weight gained from emotional over-eating, which is *Kummerspeck* (literally 'grief bacon'). English doesn't have a term and so is in the process of adopting *Kummerspeck*; hence, it hasn't yet lost its capital letter. Many of these loanwords belong to the highfalutin' level of our vocabulary – for example, *Weltanschauung* 'world view', *Schadenfreude* 'the pleasure you take in another's misfortune' and *Gestalt* 'pattern of experiences'. These still retain a fairly German-like pronunciation; many also continue to be written with the capital letter characteristic of German nouns. But if the words survive (and especially if they break out of their specialist context), they will more than likely assimilate. In English we have taken the literal *Sprachgefühl* 'language feeling' (mentioned in Table 4.4) and given it an idiomatic twist, using it now to mean the essential character of a language. It's strangely literal at the same time, though – perhaps 'language feel' is a bit like the brain's version of the 'mouth feel' of food. Even more dramatic have been the changes to the specialist term *Angst*. It first entered English in the realm of psychoanalysis to describe a kind of neurotic anxiety or dread, but has now lost its capital letter, acquired a new English vowel (although this is not reflected in the spelling) and generalised its meaning – *angst* now describes any kind of worry or fear.

Sometimes ideas can be borrowed, but not the actual words; these special types of borrowings are known as **calques**. English stole the semantics of the German *Ohrwurm* (literally 'earworm'), for example, to describe that really annoying little bit of music that rattles around inside your head, sometimes for days. This expression *earworm* shows the German idiom calqued into English. An older example is the English expression *power politics*, which is a calque on the German word *Machtpolitik*, a compound of *Macht* 'power, strength' and *Politik* 'politics'.

English borrowings into German are commonplace; they also undergo phonological (and orthographical) changes, and are usually integrated into the grammar (for example, a verb such as *chatten* 'chat' shows the German verb ending -*en*). Additional examples are provided in Table 4.9.

Table 4.9 Examples of English borrowings into German

Word	Translation	What has been borrowed
Computer	'computer'	Form and meaning
Download	'download'	Form and meaning
Happy End	'happy ending'	Form (adapted) and meaning
Handy	'mobile/cell phone'	Form
Longdrink	'mixed drink'	Form
[illegible]	[illegible]	Meaning (as a calque)
Seifenoper	'soap opera'	Meaning (as a calque)

English has two examples of very heavy lexical borrowing taking place. As we mention in Chapter 1, one relates to its contact with French. After the French-speaking Normans conquered England in

1066, the English language was well and truly overshadowed by Norman French. The prestige, or educated, language was French (and also Latin), and English was the language of the street. English eventually prevailed, and French was ousted, but the effect of French on the English lexicon was remarkable. Some 10 000 words were adopted – the majority between 1250 and 1400. It was the Normans who controlled the state, the military, and cultural and intellectual interests, and French words flooded into these areas. *Leeches* 'medical practitioners' became *doctors, physicians* and *surgeons; leechcraft* and *leechdoms* were replaced by *medicine* and *medicines* (*leech* comes from an Old English verb meaning 'to heal', and the sense 'aquatic blood-sucking worm' was a later development). The influence continued into the modern period; in the 1700s, *dentists* replaced *tooth drawers*. Who knows – had it not been for the French (and also Latin), the name of our own discipline of *linguistics* might well have been *speechlore*!

The 'Frenchification' of English continues

French-inspired vocabulary continues to dress up the language, especially in areas like food. How much classier a meal becomes when a *leek tart* is changed to *flamiche aux poireaux* and *tossed salad* to *salade composée. Broth* versus *consommé* (*potage* or *bouillabaisse*), *slice* versus *tranche, aged* versus *affiné, in aspic* versus *en gelée, reheated* versus *rechauffé* – no comparison! The language of our cookbooks and restaurants drips with these French borrowings (and the older ones, such as *soup* and *stew*, have been well and truly Englishified).

Aspects of French grammar are also used much like bits of garnish. Just the French definite article is enough to suggest fine food. Of course, it doesn't have to be the correct one – the masculine *le* is the unmarked article of menuese in dishes such as *Le Crabmeat Cocktail* and *Le Peach Melba Sundae*. Compare the use of the masculine *le* with the feminine form *petite* 'small' that commonly appears in examples like *Petite Filet Mignon* – the English word *small* never appears on menus. (In fact, you'll notice *small* is generally avoided in foodspeak; even products that come in more than one size often start at *large* and move to *extra-large, king-size* or *jumbo*.)

In their glorious account of American restaurant menus, linguists Anne and Arnold Zwicky give examples of extraordinary linguistic hybrids, such as *Stuffed Tomato aux Herbes, Shoreham Style, Café American, Broiled Steak Minute, Flaming Coffee Diablo, Prepared en Vue of Guest* and, our favourites, *Clams Larry* and

Baked Stuffed Lobster Larry. All have that truly *'je ne sais quoi'* word order of French – in the tradition of culinary greats, such as *Tournedos Rossini* and *Steak Diane*, the modifiers follow what is being modified. Names like these are a world away from some of the old-time English favourites – *mushy peas, spotted dick* (or *treacle dick*), *toad in the hole, Sussex puddle* and *bloated herring*. Not a lot of poetry there!

But the borrowings aren't always stylistic. French can often capture finer gastronomic distinctions than English. We have no term for 'a chunk of meat or fish that is longer than it is wide' or 'the edible heart of a vegetable' – the French *tronçon* and *trognon* plug these gaps. For those in the business of food, it is important to know whether the vegetables are *chiffonade* 'cut into very thin ribbon-like strips' or *julienne* 'cut into fine matchstick strips'. And there is probably a significant difference between vegetables that have been diced into a *brunoise* (tiny cubes) and those that have been diced into a *mirepoix* (not so tiny cubes).

What would the English culinary lexicon be like without this continued French input?

Contact between English and Scandinavian languages was similar, but with a different bunch of invaders and a very different outcome. At various stages during the ninth and eleventh centuries, there were Scandinavian raids on England. These culminated in 25 years of Danish rule, with large settlements of Scandinavians, particularly in the northern and eastern parts of the country, and considerable influence on native English. However, this time the borrowings were different. The expressions that English stole from Scandinavian are striking for their everyday nature. They include: basic verbs, such as *die, get* and *give*; basic nouns, such as *husband, fellow* and *egg*; and even grammatical words, such as *they, them* and *their*. These are just a handful of the hundreds of words that entered the language at this time. Borrowings of this sort suggest two things. First, unlike the French situation, Scandinavian and English existed side by side with more or less equal prestige. Second, the languages must have been very close – so close, in fact, that it is actually difficult to assess the true extent of the Scandinavian contribution. There could have been many more borrowings, but we just can't tell.

As a final example of a language that draws heavily from outside to expand its lexicon, take Japanese. It has absorbed vocabulary from many other languages; in fact, there are entire dictionaries of *gairaigo* 'loanwords'. Since the nineteenth century, English has had considerable influence on the Japanese lexicon, and in particular, *cocacolonisation* (the spread of popular American culture) has in recent years seen the flourishing of *gairaigo* in areas such as fashion, cosmetics, food, sport and music. Table 4.10 shows just some examples. As you can see, these loans have been thoroughly Japanised – in other words, adapted to fit into Japanese phonology and grammar (see Chapter 8 for more examples).

Table 4.10 English loanwords into Japanese

Fashion/cosmetics [illegible]
[illegible] (ball)’
[illegible] arine)
[illegible] club with music)
[illegible] and set’
[illegible] in baseball)

4.6 Word mortality

> Just as woods change their leaves with the onward movement of the years, the first leaves falling: even so when words advance in age, they pass away, and others born but lately, like the young, flourish and thrive.
>
> Horace (Quintus Horatius Flaccus), *Ars Poetica*, 19–18 bce

People are very aware of the appearance of new words, but tend not to notice those that are quietly dropping by the wayside. Dictionary editors, of course, have to be more aware of **lexical obsolescence**. They need to make decisions all the time about whether they classify a word as **archaic** or **obsolete**, or even whether they will bother to include it at all. It is a difficult decision. Words may no longer be relevant for modern speakers, yet people need to know them to understand earlier writing. Influential texts like the Bible, and wonderful works of literature, act something like artificial life-support systems for words that are on the way out.

There are many reasons why words disappear. When ideas, objects or institutions are no longer important to us, the words for them will drop from our mental dictionary. There is no longer much call for the verb *gymnologise* ‘to dispute naked’ (< Latin ‘naked’ + ‘speech’) – if there ever was! And in certain areas, such as clothing, lexical obsolescence is commonplace. We no longer have use for words such as *bustles* and *codpieces* (both kinds of ‘doublespeak’ clothing that draw attention to the tabooed body parts, while at the same time keeping them under wraps). Food shows a similar high turnover of vocabulary.

Clearly, societal change will be another driver of lexical obsolescence. The traditional vocabulary of sin and virtue provides us with an interesting example. As lexicographer Geoffrey Hughes has discussed, terms such as *honour, virtue, modesty, chastity* and *virginity* no longer form part of most people's active vocabulary (when did you last use, or even hear, the word *temperance*?).

Words will also disappear if they get too short. As Chapter 11 describes, sound change is generally reductive and it can happen that the mutilation is so severe that words are reduced to virtual fragments. It's hardly surprising that English swiped *law* (originally *lagu*) from the Vikings — Old English *ǣ* wasn't much of a word to describe the rule of conduct imposed by authority. Not unexpectedly, Old English *īg* 'island' and *ēa* 'river' didn't linger either. Words have to retain a certain phonetic saliency if they are to remain a useful part of the vocabulary.

Sometimes words disappear because they sound too much like other words. They might even be identical, as in the two verbs *to cleave* (originally from separate Old English verbs): *cleave* 'to slice' (compare *cleaver* and *cleft*, from the Old English *clēofan*) is now pushing out *cleave* 'to stick' (from the Old English *clīfan*). Clearly, problems are more serious when words belong to the same category and the same sphere of ideas. In German, the adjective *arm* 'poor' and the noun *Arm* 'arm' are not likely to ever be confused. Even opposite meanings can coexist quite peacefully, if they occur in such different contexts that there is no chance of a misunderstanding. *Blunt* instruments are dull and not sharp, for instance, but *blunt* comments are pointed and very sharp. When you *dust something*, it frees it from dust – that is, unless you are dusting for fingerprints or dusting a cake with icing sugar.

However, if the clash involves taboo, even separate contexts can't save words. Speakers will drop them like hot cakes if they sound too much like expressions that are offensive or embarrassing. Innocent bystanders such as *fuk* 'a sail' and *feck* 'purpose' had absolutely nothing to do with the F-word but that didn't save them. Also, witness the demise of *coney* 'rabbit' (when its pronunciation changed to rhyme with *honey* it clashed with a bawdy name for a body part).

Finally, there is simple wear and tear. Through use, novel and exciting ways of saying things become everyday and mundane. Speakers are always on the lookout for new and interesting ways to express themselves, and inevitably many fine expressions will fall by the wayside. Some areas of the lexicon are particularly prone to renewal – insults lose their wounding capacity, swear words their pungency, slang expressions their vigour. Even so, we have to admit that many apparently useful words have bitten the dust for no apparent reason – *gry* 'dirt under the fingernails' and *gound* 'sleepy dust in the eye' would seem to be useful, as would *velleity* 'mild desire but too slight to lead to any action'. We still have *gormless*, so why did we lose the noun *gorm*, meaning 'understanding'? Once people could be *kempt, corrigible, ert* and *wieldy*, and full of *ruth* and *list*, but for some curious reason these words have left us.

Further reading

The first pressing question of the chapter, 'what rhymes with orange?' is best tackled with reference to a rhyming dictionary – most of its rhyme-mates are spelled -*inge*: it's not the most poetical family of words, but includes *binge, hinge, fringe* and *syringe*. On the question of what constitutes a word, see the enduring account in Bloomfield (1933), and for more recent accounts across a wide range of languages see contributions in Dixon and Aikhenvald (2007). For more on the history of the telegram, see Hochfeder (2012); our data come from page 79. Tips on telegraph composition are from Ross (1928). More about the collocation checker for Japanese is available in Pereira, Manguilimotan and Matsumoto (2016). On presidential names, see Smith (1998). John Oliver's discussion of the onomatopoeic quality of Trump's name is available on the *Last Week Tonight* YouTube channel (https://youtu.be/DnpO_RTSNmQ); in fact, there are discussions all over the internet.

Nash and Simpson (1981) provide an impressive report of name-avoidance strategies used in central Australia. Specific treatments of word formation in English are offered by Bauer (1983), Carstairs-McCarthy (2002) and Katamba (1994), and a cross-linguistic perspective is presented by Haspelmath (2002); Benczes (2019) offers a detailed discussion of the sound play in English blends and compounds; Ghomeshi et al. (2004) is the famous 'salad-salad paper' on linguistic cloning. Various accounts of bad language can be found in Anderssen and Trudgill (1990), Allan and Burridge (2006), Hughes (1991) (on the history of English swear words) and Pinker (2007a: Chapter 7). For specialised accounts of formulaic language, see Altenberg (1990). Our quotation on the seamlessness of words in speech comes from Pinker (2007a), and we recommend Bybee (2003a) for an explanation of the interesting things that happen to frequently used words. You can see the discussion of compounds in Auslan in Johnston and Schembri (2007: 131–3). Material on names is drawn from Arnold (2012) and Lentine and Shuy (1989).

Exercises and discussion points

We focus here on consolidating the concepts from this chapter using exercises that concentrate mainly on English. Exercises on the properties of words in a more diverse range of languages appear in Chapter 5 following the introduction of some additional useful concepts.

1 English – defining words

Say whether or not you think the following forms constitute single words in English; if not, determine where you would place the word boundaries. Consider each according to the various semantic, phonological and grammatical features we have discussed in this chapter. Outline any difficulties you encounter (for example, where there might be a mismatch between these features).

[illegible]
Attorney general
Runner-up
Wannabe
Young driver risk-taking research
There’s
Won’t
Waterfall
Water-drop
Water bottle
Vegeburger

2 The longest German word

In July 2013, Germany’s longest word officially ceased to exist: Rindfleischetikettierungsüberwachungsaufgabenübertragungsgesetz

The word means ‘law for the delegation of monitoring beef labelling’, and we can break it down into the following components: *Rind-fleisch-etikettierungs-überwachungs-aufgaben-übertragungs-gesetz* ‘cow-meat-labelling-monitoring-responsibility-delegation-law’. It was repealed after the European Union had lifted a recommendation to carry out tests on healthy cattle for BSE (*bovine spongiform encephalopathy*, the technical name for ‘mad cow disease’, and itself not a bad compound).

In the *Telegraph* report of the news piece (27 July 2013), the journalist observed:

> In theory, a German word can be infinitely long. Unlike in English, an extra concept can simply be added to the existing word indefinitely. Such extended words are sometimes known as *Bandwurmwörter* – ‘tapeworm words’. In an essay on the German language, Mark Twain (1880) observed: ‘Some German words are so long that they have a perspective.’

Is the journalist correct or can there ever be ‘the longest German word’? And what about English? This piece suggests English words have a finite length.

3 Word formation – *leximania*

The comedian Rich Hall has his *sniglets* (‘words that don’t appear in any dictionary but should’): *stroodle* ‘the annoying piece of cheese that connects you to your hot pizza’; *glarpo* ‘the juncture between ear and skull where pencils and pens are stored’; *sloovers* ‘the remnants of soap too small to use, but too big to throw away’; and there are many more. Make up your own sniglets (in any language) using any of the word-formation processes discussed in this chapter. Explain the words, their meanings and the process used.

4 Any language – borrowing

Take a language that has a considerable number of loan words from other languages (for example, Japanese loans from Chinese and English; Indonesian loans from Arabic, Sanskrit or Dutch; English loans from French, Dutch, German, Latin and Greek). Make a list of up to 50 loan words. Study how foreign words are adapted to the native linguistic system (phonological, grammatical and semantic). If you can, indicate what information the loans provide us about the nature of contact between the two languages.

5 English – linguistic fossils

When linguistic changes occur, they often leave behind some sort of trace – a relic of the original set-up. Words don’t just disappear; there is typically something left over. This is what prompted the Dutch linguist Van der Tuuk to once describe language as ‘something of a ruin’. Modern languages are full of relic forms that we can use to reconstruct features of vocabulary (even sounds and grammar) from earlier times.

The following words are historically compounds. Identify the elements from which each word was formed (*note:* some elements may now be obscure):

(a) werewolf

(b) cobweb

(c) gossip

(d) hatred

(e) midriff

(f) garlic

(g) lukewarm

(h) daisy

(i) window.

Research project

Word-formation processes in English

In 1991, John Algeo completed a study of new words over a 50-year period (1941–91), sourcing his neologisms from the collection 'Among the New Words' that appears each year in *American Speech*, the journal of the American Dialect Society. For each of the major types of word-formation processes, he found the percentage of new words shown below.

Algeo's study shows affixation and compounding way out ahead, and blending and borrowing to be insignificant sources for new words. Your task is to determine whether this is still the case. You could use the same source as Algeo (the 'Among the New Words' publications), or you could select another means of collecting neologisms. You will probably have access to the *Oxford English Dictionary* online through your library, and on the company's homepage under 'What's New' you can find links to full lists of new words for different years. There will also be dictionaries of new words in the library.

Type	**%**
Compounding	40
Affixation	28
Shifting	17
Shortening	8
Blending	5
Borrowing	2

For more activities, please go to www.cambridge.edu.au/academic/linguistics.

5

Morphology: The structure of words

◆

Polonius: What do you read, my lord?

Hamlet: Words, words, words.

William Shakespeare (1600–01), *Hamlet*, Act II, Scene 2

5.0 Introduction

We established in Chapter 4 that words are recognisable units of structure, but not workable units for linguistic description. One difficulty was the imprecision of the actual term *word*. There is often a problem when terms from ordinary language are used in a technical context. We end up having to redefine them and create new terms to capture distinctions that aren't covered in their everyday usage, and this is exactly what happens here.

As units of description, words have one important drawback: they aren't the smallest components of meaning (recall the problem we had with *unfamiliar* versus *not familiar*). Although words are clearly meaningful and can be used to form larger structures, you may have noticed from some of the earlier discussion that they aren't necessarily simple unitary things. In fact, in many languages there are elements within words that are used over and over again. This chapter is all about the ways that words are put together, or what is known as **morphology** (from the Greek *morphe*, 'form, shape'). We will show you how to break down words into their smallest meaningful parts. We'll also explore the ways in which these smaller units, called **morphemes**, are combined and recombined to make larger and larger structures – words of all different kinds.

5.1 Words, lexemes and morphemes

We start by looking at the internal parts of words. The morpheme is best described as the smallest unit of meaning in the structure of the language. By this we mean that we cannot divide the unit any more without severely altering the meaning. For example, *cardigan* is a word, and it is also a morpheme. As one complete unit by itself, it has meaning, and while we can apparently divide it into smaller units, such as *car, dig* and *an*, these have meanings that are not associated with the word *cardigan*. On the other hand, a word like *cardigans* has a meaning as a whole, but we can also divide it into two smaller meaningful units: *cardigan* + the plural ending -*s*.

If you take a look at the following examples, you will see that there are some elements that appear again and again in English. Which morphemes can you identify in the set of words in example 5.1?

(5.1)

- beautiful
- awful
- handful
- resourceful
- dutiful
- remorseful
- handy
- awfully
- awfulness
- remorsefulness
- resourcefulness
- handiness

Many of the words in example 5.1 share the morpheme -*ful*, used to change words referring to a quality, such as *beauty*, into words that describe something with that property, *a beautiful day*. As English speakers, we do this type of analysis even if we don't recognise all the parts of the words. You can probably identify quite a few morphemes in the words in example 5.2:

(5.2)

- discombobulate
- chriography
- decreolisation

On the other hand, sometimes surprisingly long words cannot be divided up into smaller parts (example 5.3):

(5.3)

- hippopotamus
- crocodile
- banana

Here are some examples from Mali. Can you tell what is being changed when the additional morpheme is added to each stem?

(5.4)

- *amor* ‘large’, *mamor* ‘intensely’
- *amēr* ‘good’, *mamēr* ‘properly’
- *aulul* ‘long’, *maulul* ‘a long time’

(5.5)

- *dong* ‘turn something’, *dong mes* ‘turn around’
- *par dēm* ‘drag something’, *pardēm mes* ‘crawl’
- *sek* ‘lift something’, *sek mes* ‘get up’

- *tu thēm* ‘decorate something’, *tu thēm mes* ‘decorate oneself’

In the first case, the morpheme *ma-* is added to the start of a word in Mali to change its function from describing something, *amēr* ‘good’, to reporting how something is being done, *mamēr* ‘properly’. In the second case, the morpheme *mes* is used to refer to an action that you apply to yourself, *dong mes* ‘turn around’. Without *mes*, the word refers to actions done to others – for example, *dong* ‘turn something’.

As we have seen, words come in all sorts of shapes and sizes, so another useful preliminary distinction to make is between words and **lexemes**. The label lexeme refers to a word in all its different forms. This is a really handy concept because it allows us to recognise that *cat* and *cats* or *laugh, laughing* and *laughed* are different forms of the same word. Lexemes are often written in small capital letters, with the word in its simplest form: cat and laugh. This distinction is particularly useful in languages with lots of different forms for individual words. Consider Table 5.1, which shows examples from Russian for the verb ‘take’. This is only a selection of forms for the lexeme. The complete set runs to 46 different forms! (Don’t worry about the grammatical labelling here – we will return to this later.)

Table 5.1 Imperfective forms of the verb ‘take’ in Russian (present and future tenses only)

Person	**Present tense**	**Future tense imperfective aspect**
1st person singular – *ja*	[illegible]	[illegible]
2nd person singular – *ty*	[illegible]	[illegible]
3rd person singular – *on*	[illegible]	[illegible]
1st person plural – *my*	[illegible]	[illegible]
2nd person plural – *vy*	[illegible]	[illegible]
3rd person plural – *oni*	[illegible]	[illegible]

So, the words that appear in dictionaries are actually lists of lexemes (the word forms *cat* and *cats* don’t have separate dictionary entries). Our everyday understanding of words is more concrete: words are actual assemblages of letters (or sounds), and this is what you would count if you were told to write an essay of 2000 words – it is highly unlikely that you would count lexemes!

5.2 Roots, stems and affixes

All the parts of words are known as morphemes, but they come in a range of types, depending on their function and status within a word. There are two basic types of morphemes: **roots** and the **affixes** that attach to them.

The foundation of a complex word is the root. All the words in example 5.6 contain the root *true*:

(5.6)

- truer
- truest
- truly
- untrue
- truth
- truthful
- truthfully
- untruthfully
- untruthfulness

In large words, roots are built up into more complex forms in stages. The intermediate stages are called **stems**. In a word containing only two morphemes, the root is identical to the stem. In practice, when we talk about adding morphemes onto a word, the thing we are adding to is a stem. All the roots in Table 5.2 can also function as stems or as words in their own right.

Table 5.2 Roots and stems in a selection of English words

[illegible]
Combined with [illegible]
[illegible]
Combined with [illegible]
[illegible]

Languages vary with regard to how many morphemes are available for combining into one word. English is an example of a language with only moderate processes of affixation, but in some languages all sorts of ideas are conveyed at the level of the word. Recall the example in Chapter 4 from Yupik (spoken in Alaska) – as you can see, whole sentences can be constructed from one root morpheme. In this case, it’s the morpheme *angya* ‘boat’:

(5.7)

angyaghllangyugtuq

angya	*ghlla*	*ngyug*	*tuq*
boat	big	acquire	[illegible]

‘he wants to acquire a big boat’

Any form that is attached to a stem is called an affix. There are several types of affixes, classified according to where they appear on the stem: **prefixes, suffixes, circumfixes** and **infixes**.

Prefixes are morphemes that come before the stem. Table 5.2 above demonstrates the following prefixes from English: *re-* in *recycle* and *un-* in *unbelievable*. Can you now identify prefixes in any of the Japanese words in example 5.8?

(5.8)

massugu	‘straight ahead’
masshiro	‘pure white’
maikai	‘each time’
mainichi	‘every day’
zenshō	‘burnt to the ground’
zentai	‘entire, whole’

You probably found the prefixes *ma(s)-, mai-* and *zen-*. Take a look at the roots for these words as shown in example 5.9. Knowing the meanings of the roots, how could you describe the meanings of the prefixes?

(5.9)

sugu ‘directly’

shiro ‘white’

kai ‘occurrence’

niche ‘every day’

shō ‘baked’

bu ‘partition, class’

We also find the words *ma-hiru* ‘in the middle of the day’ and *mak-kura* ‘pitch dark’. Based on these additional examples, how would you analyse the form of *ma(s)*?

Any morpheme that comes after the stem is a suffix. In Table 5.2 on p. 116, the following English suffixes were shown: *-ation* and *-al* in *derivational, -s* in *recycles, -ing* in *recycling* and *-able* in *believable*. For comparison, a range of German suffixes are listed in Table 5.3.

Table 5.3 Some suffixed words in German (based on the verb stem *ehr-* ‘honour’)

Word	**Suffixes**	**Translation**
ehren	*-en*	‘to honour’
Ehre	*-e*	‘honour’
Ehren	*-en*	‘honours’
Ehrlichkeit	*-lich -keit*	‘honesty’
ehrenhaft	*-en -haft*	‘honourable’

Words with circumfixes look like they have both a prefix and a suffix. The difference is that circumfixes involve two segments working together to encode a single meaning. Circumfixes are less common than prefixes and suffixes, but our case-study languages give us some examples. German uses the circumfix *ge-* + *t* or *ge-* + *-en* to form what is known as the past participle of the verb. (English once had the same circumfix: either *ge* + *od* as in *geplegod* ‘played’ or *ge* + *en* as in *geflogen* ‘flown’.)

(5.10)

ich	*habe*	*geehrt*		
ich	*habe*	*ge-*	*ehr*	*-t*
I	have	past participle	honour	past participle
‘I have honoured’				

Infixes are even less common than circumfixes. They occur somewhere in the middle of the stem. Homer Simpson provides one of the rare examples for English with his *-ma-* infix in words such as *edu-ma-cation, sophisti-ma-cated* and *vio-ma-lin*. Probably based on the inspiration of *whatchamacallit* and *thingamabob*, 'Homeric infixation' seems to suggest that Homer is not entirely sure what he's talking about! There is also a fine tradition of inserting expletives into the middle of words to express strong feeling (*fan-bloody-tastic*). Another intensifying infix is illustrated by Roald Dahl's *scrum-diddly-umptious* (here *diddly* is infixed and the rhyme *um* is repeated). In fact, English has a long history of having fun with infixes – a very early example is *-de-* (originally mock French) in *flibberdegibbet* 'evil spirit', *grizzledemundy* 'grinning stupid person' and *slubberdegullion* 'miserable worthless wretch'.

Some languages have infixes that are more than simply playful aspects of the lexicon, as in English. Within the Austronesian language family, you will find many examples of regular infixing, and these affixes integrate more closely with the grammar. Tagalog (of the Philippines) forms active verbs with an *-um-* infix (see Table 5.4).

Table 5.4 Infixes in Tagalog

Stem	**Meaning**	**With infix**
sulat	'to write'	*sumulat*
ʔabot	'to reach for'	*ʔumabot*
gradwet	'to graduate'	*grumadwet/gumradwet*
preno	'to brake'	*prumeno/pumreno*

Affixes (prefixes, suffixes, circumfixes and infixes) are always **bound**. This means that they cannot be used as words on their own – that is, as **free morphemes**. Most English roots are free; under the right circumstances, they can occur on their own as words without any additional morphology. But English also has roots that are bound. Often these have been borrowed from other languages (such as Greek or Latin). Examples include the root *-lingual* in words such as *bi-lingual-ism* and *mono-lingual*.

Some of these bound roots contribute to large numbers of English words. This is true of *-ceive*, which is at the base of *conceive, deceive, misconceive, misperceive, perceive, preconceive, receive* and many more. These words don't show a constant meaning in their root, but there is a core sense there, which is then modified by the affix. Often the meaning of the root can only be determined historically. The root *-ceive* 'take' is from Latin, and although it is no longer meaningful for modern speakers, there are good reasons to recognise it as a root. Segments such as *per-, re-* and *con-* occur with great frequency and are undoubtedly prefixes. Furthermore, the root behaves in the same way with respect to derived forms.

(5.11)

perceive, perception, perceptive
receive, reception, receptive
conceive, conception, conceptive

Only if we recognise *-ceive* as a morpheme can we give a uniform account of these derived forms. So, while the semantic basis for recognising these morphemes remains shaky, these other pieces of evidence prompt us to adopt this analysis. Such examples are problematic for our earlier account of morphemes as minimal meaningful units, since we are identifying them not so much on the basis of their meaning (which we had trouble pinning down) but more on their distributional behaviour. One linguist, Francis Katamba, describes these root morphemes as more 'chameleon-like' in their meanings – they change according to their environments, just as chameleons will change their skin coloration.

Crans and *huckles* – more bound morphemes

Consider the following list of berry words and decide where you might put the morpheme boundaries: *blackberry, blueberry, cranberry, boysenberry, loganberry, gooseberry, strawberry, raspberry*.

Presumably, you want to identify a *berry* morpheme here. There is no problem in *blackberry* and *blueberry*, but what do you do with *boysen, cran* and *logan*, which in isolation have no meaning? These are examples of special bound morphemes (appropriately dubbed 'cranberry morphemes' or 'cranmorphs'). History can usually shed some light. *Cranberry* derives from the name for the bird, though no one is entirely sure why this fruit is linked to cranes. Less of a problem is *boysen*. True, it's also only meaningful in combination with *berry*, yet its link to the American horticulturist Rudolph Boysen is clear – he developed this hybrid in the 1920s. Similarly, it was Mr Logan who inspired the *loganberry*.

At first blush, *gooseberry* doesn't look problematic, since it appears to be a straightforward compound of *goose* + *berry*. The question is why *goose*? Are these birds great consumers of berries, or did people choose the name because the rough surface of the berry reminded them of prickly gooseflesh? Researchers have uncovered earlier versions of the word: *groser* and *gozell*. These don't look much like *goose* at all and probably came from the French word for the berry *groseille*. English speakers are well known for remodelling foreign imports to look like English words and must have changed *groseille berry* into the more linguistically friendly *gooseberry* – the justification would have been the bristly goose-bump textured surface of the berry.

So, what about the *straw* in the *strawberry*? Straw is the name we give to the stalks of threshed crops, something we once used to strew on the floor (and the word is related to *strew* 'to spread or scatter'). One suggestion then for the origin of *strawberry* is that the long trailing runners of the strawberry plant reminded people of pieces of straw on the floor. You could think of the word as 'strewnberry'. Another possibility is that it is a remodelled form of *strayberry* – after all, the runners do stray. Straw was also once used to protect strawberries, yet the wild fruit was called *strawberry* well before it was propagated in gardens. Then again, wild strawberries grow in grassy places and in hay fields, and their external seeds look like straw. There seems no end to the possible suggestions! In fact *strawberry* may well be the result of a combination of these ideas – a kind of lexical bitser, with multiple origins.

The difficulty is deciding how much history should inform our morpheme boundaries. Consider the *raspberry*. Those of you who have picked raspberries and been scratched by their thorns are probably tempted to make a connection with the verb *rasp* 'to grate', and would analyse this as a straightforward compound. However, no connection with *rasp* has ever been established. This task of carving up berry words into morphemes proves difficult. And we haven't even started on *bilberries, huckle-berries, hurtleberries* and *whortleberries*! Most berry words aren't nearly as obliging as the *blackberry* or *blueberry*.

Along the continuum from free to bound morphology, somewhere between an independent word and an affix lies a class of morphemes known as **clitics** (the root here is the Greek verb meaning 'to lean'). In speech, clitics 'lean on' neighbouring words rather than being pronounced as independent words in their own right. English clitics generally represent words that can also function fully independently and, in writing, some are indicated by an apostrophe, often with part of the equivalent independent word omitted. Examples are given in Table 5.5.

Table 5.5 Clitics in English

Cliticised form	**Host**	**Independent form**
can't	can	not
she's	she	is
Sarah'd	Sarah	had

You can see that these clitics are bound morphemes in the sense that they attach to other words, but they are not affixes, in that they correspond to elements that can be expressed as independent words. When we look at the creation of grammar in Chapter 11, we will see that over time words can develop into affixes through cliticisation: the development *not* > *-n't* shows this transition from full word to clitic to (potential) affix.

5.3 Inflection and derivation

Affixes can have two very different types of functions. The first is to give grammatical information about the stem. This is called **inflectional morphology**. For example, English verbs take the following suffixes to give information about the time an event or action takes place (this involves categories called **tense** and **aspect**, which we examine in a little while):

Table 5.6 Inflectional forms of the verb *dance* in English

Suffix	**Inflected word**	[illegible]	**Example**
[illegible]	dances	Third person singular present tense	*She dances with heaps of enthusiasm.*
[illegible]	danced	Past tense (and past participle)	*He danced the whole night long.*
[illegible]	dancing	Progressive	*They* [illegible] ***dancing*** *the night away.*

The other type of morphology is called **derivational morphology**. In this case, morphology changes the meaning or function of the stem, making (or deriving) a new word (or lexeme). Again using the stem *dance* in English, we can apply derivational suffixes to create new words.

Table 5.7 Derived words based on the verb *dance* in English

Suffix	**New word**	[illegible]	**Example**
[illegible]	[illegible]	Change verb to noun	[illegible] *are inspired to* ***dance*** [illegible]

dance	**danceable**	adjective	*That's a really **danceable** tune.*
dance	**dancing**	noun	*Dancing is great for your health.*

Similarly, in Māori, the prefix *kai-* is available with a range of stems to make a new word that refers to a person who does the action encoded in the stem.

Table 5.8 Derived words using the prefix *kai-* in Māori

Stem	**Meaning**	**New word**	**Meaning**
ako	'teach'	*kaiako*	'teacher'
kōrero	'speak, story'	*kaikōrero*	'speaker, story-teller'
mahi	'work'	*kaimahi*	'worker'
titiro	'look'	*kaititiro*	'spectator'
waiata	'sing'	*kaiwaiata*	'singer'
whakahaere	'administer'	*kaiwhakahaere*	'administrator'

Derivational affixes are used in word formation; they create new lexemes and therefore belong in dictionaries. Inflectional affixes create new word forms of lexemes and belong more in grammar books. For instance, the verb *judge* brings to mind the action of adjudicating or mediating. We can convert this into a noun by adding a *-ment* suffix, and what then comes to mind is the product of that action. However, if we add an inflectional suffix to *judge*, such as *-s* or *-ed* (as in *he judges/judged*), we have the same concept of 'adjudicate' (the action). The forms *judge, judges, judged* and *judging* are simply different forms of the one word (or more correctly the lexeme), JUDGE; JUDGEMENT is a different lexeme altogether, with its own set of forms.

It can sometimes be difficult to tell whether or not a complex word involves derivational or inflectional affixes. Laurie Bauer offers a useful test that involves checking whether or not the word can be replaced by a single morpheme. Here's an English example to illustrate:

(5.12)

(a) The **runner** fell and broke his leg. (*runner* involves derivational morphology)

(b) The **man** fell and broke his leg.

(5.13)

(a) He **runs** extremely well. (*runs* involves inflectional morphology)

(b) He ***sprint** extremely well.

The inflected form in example 5.13a cannot be replaced by a monomorphemic word like *sprint*, so we've used an asterisk to indicate that the sentence is unacceptable. Because *sprint* can't replace *runs*, we can conclude that the suffix -*s* is inflectional.

tlhIngan Hol (Klingon) morphology

We saw in Chapter 2 that trained linguists are sometimes called upon to construct languages ('fantasy conlangs') for humanoids and other alien types (the Na'vi language in *Avatar*; Klingon in *Star Trek*) and a range of mythical creatures (the Elvish language in the *Lord of the Rings* trilogy; Parseltongue in the *Harry Potter* series). Many of these languages have since been expanded by their creators, or perhaps fans, to become learnable languages. Grammars have been published and for some of these fantasy languages, such as Na'vi and Klingon, tutorials are available on the internet. There is even a Klingon version of *Hamlet*.

Klingon morphology is notable for its complex system of noun and verb affixes. New nouns take a range of derivational affixes to which different inflectional suffixes may then be added. Something we haven't mentioned yet is ordering of affixes – like words, these are always stacked in a definite ordering relationship (as speakers of English, you would not accept **preunmeditated* or **judges-ment*!). In Klingon grammar, noun affixes are numbered for convenience. The noun stem 1-2-3-4-5 indicates the five types of affixes:

- *type 1:* augmentative/diminutive (indicating relative importance or power of a noun)
- *type 2:* number (an optional suffix indicating plural)
- *type 3:* qualification (indicating the speaker's attitude)
- *type 4:* possession (indicating possessor)
- *type 5:* grammatical relations (indicating the function of a noun in the sentence).

Here is a noun with five suffixes: *QaghHommeyHeylIjmo’* ‘because of your apparent minor mistakes’. The following is its morphological breakdown:

Qagh-Hom-mey-Hey-lIj-mo’

mistake-diminutive-plural-apparent-your-because of

Verb stems have a prefix indicating the players in the event or action (so they correspond to free-standing pronouns in English, such as *I, me* and *he*). The slots in the Klingon verb look like this: prefix-verb stem 1-2-3-4-5-6-7-8-9. The suffixes express various concepts that are important to Klingon culture (such as accuracy and social hierarchy). The language doesn’t have tense, but many of these suffixes indicate other kinds of temporal information – for example, they indicate the stage of an event, action or state, and how it relates to the flow of time. In between suffixes 1–9 appear special markers called ‘rovers’. One such roving suffix negates the verb.

Paradigms

We have seen that affixes that encode grammatical information are called inflectional morphemes. In languages that make a good deal of use of inflectional morphology, there are large sets of inflected forms for each word, which are called **paradigms**. We touched on patterns and paradigms in Chapter 1, but here we explore them in more detail.

Table 5.9 shows part of the inflectional paradigm for the noun *boy* in Russian. Such paradigms are not just handy ways of setting out the different forms that words can assume in languages like Russian. As we will see in later chapters, they are the territory in which changes can take place; they also seem to have psychological reality for speakers and foreign language teaching makes use of this.

Table 5.9 Partial inflectional paradigm for BOY in Russian

Case	**Singular**	**Example (singular)**
Nominative	*mal’čik-Ø*	*Mal’čik pročital knigu.* ‘the boy read the book.’
Genitive	*mal’čik-a*	*Kniga mal’čika* ‘the boy’s book’
Accusative	*mal’čik-a*	*Ja videl mal’čika.* ‘I saw the boy.’

The rows in Table 5.9 show various forms for the noun (note that Ø signals that there is no ending). The different endings have to do with **case**. Case relates to the role of the noun in the sentence. It is no longer an important inflectional category in Modern English (pronoun forms such as *he/him* and *who/whom* are relics), but it remains important in most European languages. The other value in the table is **number**. In English, there is a contrast between singular and plural. The singular is unmarked while the plural typically takes the suffix *-s* (*cat* versus *cats*). Russian also has a contrast between singular and plural.

You might be wondering why we have made the distinction between genitive and accusative case here, given that the forms are identical. The reason is that the difference between these cases is apparent in other parts of the system. Compare Table 5.9 showing the word *boy* with Table 5.10 showing the word *orange*. The nouns *boy* and *orange* belong to different **gender** classes in Russian. *Boy* is masculine while *orange* is feminine. There is also a neuter class in Russian, containing words such as *window* and *name.* The label ‘gender’ is a confusing term (it’s hard to be sure how oranges should be feminine), but if you forget the sex associations and concentrate on its original meaning of ‘class’ or ‘kind’ (Latin *genus*), its grammatical application becomes clearer (and bear in mind that languages can have up to 20 ‘genders’ – Mali has nine!). Many languages divide nouns into different classes (at least partly) based on semantic and phonological criteria, though time often obscures these.

Table 5.10 Partial inflectional paradigm for ORANGE in Russian

Case	[illegible]	Example (plural)
Nominative	*apel’siny*	*Apel’siny zdorovye*. ‘oranges are healthy.’
Genitive	*apel’sinov*	[illegible] *apel’sinov* ‘the colour of the oranges’
Accusative	*apel’siny*	[illegible] *apel’siny*. ‘I ate the oranges.’

In the case of ‘orange’, the words in the rows for nominative and accusative are the same. Inflectional paradigms often have interesting patterns such as this, and linguists can learn quite a lot about the history of a language by looking closely at these patterns. This is something we return to in Chapter 11.

Case-marking cruelty?

Many languages classify nouns – some very elaborately. Like Russian, German has a gender system (*der Hund* ‘dog’ is masculine, *die Katze* ‘cat’ is feminine and *das Mädchen* ‘girl’ is neuter). For a bit of fun, we want to quote a piece where American novelist Mark Twain hilariously explains what happens to dogs and cats once they enter the gender and case system of German.

> A dog is ‘*der Hund*’; a woman is *‘die Frau’*; a horse is ‘*das Pferd*’; now you put that dog in the genitive case, and is he the same dog he was before? No, sir; he is ‘*des Hundes*’; put him in the dative case and what is he? Why, he is ‘*dem Hund*.’ Now you snatch him into the accusative case and how is it with him? Why, he is ‘*den Hunden*.’ But suppose he happens to be twins and you have to pluralize him – what then? Why, they’ll swat that twin dog around through the 4 cases until he’ll think he’s an entire international dog-show all in his own person. I don’t like dogs, but I wouldn’t treat a dog like that – I wouldn’t even treat a borrowed dog that way. Well, it’s just the same with a cat. They start her in at the nominative singular in good health and fair to look upon, and they sweat her through all the 4 cases and the 16 the’s and when she limps out through the accusative plural you wouldn’t recognize her for the same being. Yes, sir, once the German language gets hold of a cat, it’s goodbye cat. That’s about the amount of it.

Changing words such as ‘dog’, ‘cat’ and ‘the’ according to the grammatical system Mark Twain describes here is called *declining* a word. In his 1880 essay ‘The Awful German Language’, Twain concludes, ‘I’d rather decline two German beers than one German noun!’ (Apparently Twain’s unsuccessful attempts at learning German began at age 15.)

A thousand years ago, English nouns also fell into one of three classes: masculine, feminine and neuter. True, these might have roughly corresponded to the sex or sexlessness of objects, but not necessarily (Old English *wifmann* ‘woman’ was masculine and *mædencild* ‘girl’ was neuter). English has now lost this system of grammatical gender. Thus, although the Russian gender system applies across all nouns, in English usually only human nouns and some domestic animals are treated as having gender. All other nouns are neuter. This is called biological gender and even then this only comes into play with the pronouns *he, she* or *it*. Sometimes this type of gender system is called ‘natural gender’. Mind you, English speakers often assign a gender to objects with which they are emotionally involved (Kate’s hurdy-gurdy is definitely a ‘he’ and Tonya’s car is a ‘she’).

Another interesting way that inflectional paradigms can vary involves changes to the stem. The English verb *be* is a good example. It has a number of different forms depending on how it is being used (e.g. *am, is, are, was, were*, etc.). In addition to number, the other important values involved are **person, tense** and **aspect**. Let's take a moment to look at these concepts before we consider how they relate to *be*.

Languages distinguish between different person values associated with the speaker, the hearer and others who are not in the conversation. These contrasts are called first person (the speaker or writer), second person (the hearer or reader) and third person (anyone or anything else). This is apparent in the English pronoun system, as shown in Table 5.11.

Table 5.11 The English subject pronoun paradigm

Person	**Singular**	**Plural**	
First	*I*		Speaker
Second	*you*		Hearer
Third – masculine	*he*	*they*	Other
Third – feminine	*she*		
Third – neuter	*it*		

Tense and aspect are grammatical categories that give information about when an event happened and how this event relates to the flow of time. Tense allows us to place the event in time while aspect describes the way the event is characterised by the speaker. For example, depending on the language, an event may be characterised as complete or incomplete, a one-off event or an event that recurs, an extended event, or an event that occurs habitually. Note that the perfect aspect is also called 'completed', as it refers to events that are seen as complete. Progressive aspect is also called 'continuous', as it refers to events that are being viewed from within. Consider example 5.14.

(5.14)

past perfect	*She had enroll**ed** herself at the university.*
present perfect	*She has enroll**ed** herself at the university.*
past (no aspect marked)	*She enroll**ed** herself at the university.*
present (no aspect marked)	*She enrols herself at the university.*
past progressive	*She was enroll**ing** herself at the university.*
present progressive	*She is enroll**ing** herself at the university.*

These four concepts, the person and number of the subject as well as tense and aspect, are all important for the formation of verbs in

English. All of them are represented in Table 5.12, which shows the different forms of the verb *be*. Note how the stem *be* only appears in less than half of the forms; later in the chapter, we will explain why this is.

Table 5.12 Inflectional forms of the verb BE in English

Verb form	**Example**
Infinitive	*Could you please* ***be*** *quiet?!*
First person singular present tense	*I* ***am*** *so pleased to meet you.*
Third person singular present tense	*She* ***is*** *such a dedicated student.*
General present tense	*You* ***are*** *a great class to teach.*
Progressive	*Dan is* ***being*** *funny!*
First and third person singular	*He* ***was*** *delightful.*
General past tense	*They* ***were*** *an interesting bunch.*
Past participle	*Where have you* ***been****?*

Even though the forms vary greatly, they are all inflections of the verb *be*. By studying these forms, and referring back to texts in earlier versions of English, as well as looking at related languages, we have evidence of the inflectional richness that once existed in English, and we give you a glimpse of this in Chapter 11.

5.4 Morphs and allomorphs

Earlier we described the morpheme as the minimal unit of meaning – for instance, *cats* consists of two morphemes (*cat* + *s*). Here the meaning is 'feline quadraped' + 'plural', but the physical forms (in other words, the actual speech sounds to which the meanings attach) are called **morphs**. Going back to the concept of the **sign** from Chapter 1, we can say that the morph is the form of the sign and it signifies (denotes) a particular meaning.

More often than not, morphemes have more than one form. These various forms are called **allomorphs** (from the Greek *allos* 'other' + *morphe* 'form'). Allomorphs are simply variants or alternative pronunciations (phonetic forms) of morphemes. As it turns out, *cat* (or, in phonetics, [kʰaet]) is the only allomorph of 'feline quadraped'. But there are many examples where morphemes have a number of different shapes; that is, they have a number of allomorphs. A good example is the English indefinite article. In this case, English has two allomorphs: *a* and *an*.

(5.15)

a beer	**an** iced tea
a main dish	**an** entrée
a starter	**an** antipasto
a Lancashire hotpot	**an** Irish stew
a poached egg	**an** omelette

Allomorphs typically appear because there is some sort of conditioning factor involved; they are triggered by something in their environment. If you examine the two columns of words in example 5.15, you will see that words beginning with a consonant produce the allomorph *a*, and words beginning with a vowel produce *an*. These allomorphs are said to be in **complementary distribution** because they occur in completely different environments. (We'll put aside the thorny issue of whether you say *a historical fact* or *an historical fact*!)

In this case, the allomorphy is reflected in the spelling, but this is not always the case. Turn the forms in example 5.16 into the plural and think about how you pronounce them.

(5.16)

- cat, hat, sack
- dog, hog, bag
- price, creche, cheese

The plural for each of these words is written as 's', but in fact this has three different pronunciations: the first is a voiceless sound, [s]; the second is a voiced sound, [z]; and the last has a vowel before it, as in [əz]. So, the plural suffix in English has three alternative realisations, or allomorphs. (The upside down *e* [ə] is called a 'schwa'; we'll explain how it is used in Chapters 7 and 8.) The distribution of these different allomorphs depends on the sound of the final segment of the stem, and when new nouns such as *grotch, wug* and *sputnik* come into the language, native speakers of English select between these alternative forms quite subconsciously and automatically assign the correct plural. This allomorphy is **phonologically conditioned**, meaning that the allomorphs reflect relationships with neighbouring sounds.

It isn't just affixes that come in different forms. Strange things can also happen to the pronunciation of roots and stems, and we sometimes have competing forms (or allomorphs) here, too. Since we are addressing the plural in English, consider the additional data shown in example 5.17.

(5.17)

- scarfscarves
- loafloaves
- wifewives
- househouses

Here you can see the variation in the form of the stems; it is reflected in the spelling for the first three, but not for *house* (we can show it phonetically: [haus] + [plural] → [hauzəz]. (We are aware that this is a changing aspect of English, so don't be concerned if you don't have different stems for the singular and plural forms of these words – this is an irregularity of the modern language that some speakers are ironing out!)

These examples involve allomorphy, as we saw earlier, except that this time they involve the stem. Specifically, there is a difference in voicing in the last consonant of the stem – in the examples here, a voiceless fricative ([f] or [s]) becomes a voiced fricative ([v] or [z]). Don't be put off too much by these phonetic details – they will be covered in Chapters 7 and 8. For the moment, all you need to absorb is that this pattern is not general in English. Although there are many words that exhibit this allomorphy (for example, *elf, wreath, bath, life* and *mouth*), there are many more nouns (and increasingly so) that do not change to a final voiced consonant when the plural is added (such as *cliffs, places* and *sloths*). To understand this alternation in the stems, we have to go into the history of English, which we aren't able to do here. From the viewpoint of Modern English, all we can say is that there is nothing in the phonological make-up of these words that is causing the alternation. They are special facts about English that we have to learn along with the meaning of the words themselves. This allomorphy is lexically conditioned.

Now consider one last batch of nouns in English:

(5.18)

- sheep, deer, fowl
- tooth, mouse, man
- ox, child, criterion

These nouns are all badly behaved when it comes to forming plurals. The first group has a zero-morph; in other words, the words *sheep, deer* and *fowl* remain unchanged in the plural. The nouns in the second group show a change in the vowel: ‘tooth’ + ‘plural’ = *teeth*. The last group is a motley set of irregular plurals: *oxen, children* and *criteria*. Like the groups in example 5.17, there is nothing about the shape of any of these nouns that triggers these unusual plural forms; the distribution of the plural allomorphs can only be stated in relation to the list of morphemes – language learners simply have to memorise the forms. These are also lexically conditioned.

As a final example of allomorphy, consider the variation that German shows in the stems of some of its nouns. As in the English *wife–wives*, certain final sounds of noun stems are voiceless in the singular, but voiced in the plural. This is not reflected in the spelling of Modern German, as shown in Table 5.13. We have provided phonetics alongside the spelling so you can see what is happening here (some phonetic symbols will appear strange, but just focus on the final consonants of the stems).

Table 5.13 Allomorphy in German noun stems

Singular	**Plural**	**Translation**
Tag [taːk]	*Tage* [taːgə]	‘day’
Hund [hunt]	*Hunde* [hundə]	‘dog’
Los [loːs]	*Lose* [loːzə]	‘lot (for deciding)’
Nerv [nɛrf]	*Nerven* [nɛrvən]	‘nerve’
Kalb [kalp]	*Kälber* [kɛlbə]	‘calf’

We will revisit these German examples in Chapter 11 when we look at other examples of phonologically conditioned allomorphs.

5.5 Lexical and grammatical morphemes

> If the English language had been properly organized … then there would be a word which meant both ‘he’ and ‘she’, and I could write, ‘If John or Mary comes, heesh will want to play tennis,’ which would save a lot of trouble.
>
> A.A. Milne (1930), *The Christopher Robin Birthday Book*

Recall that in Chapter 4 we made a distinction between lexical words and grammatical words based on the content and function of the words. Lexical words are words that:

- have clearer and more tangible meanings
- belong to large open classes (nouns, verbs, adjectives, adverbs)
- may be readily expanded through borrowing and derivation.

Grammatical words are words that:

- have highly abstract meanings
- belong to small closed classes (pronouns, conjunctions, prepositions)
- are not usually added to the language.

Now we've introduced the concept of the morpheme, it makes more sense to talk about lexical and grammatical morphemes, since this allows us to include the various types of affixes and other bound morphemes we looked at earlier. To demonstrate the difference between lexical and grammatical meaning, consider the following two short paragraphs.

(5.19) Crip osh plam fi hap simplix entin flimpt delicious saladem op eesh servist quern tomatoem oshin ok season. Slice hap tomatoem okhap small piecem. Squash heom trog epin sieve. Chop hap garlic crosh finekin crosh heesh kennte entin add sheesh op hap tomatoem. Mix well.

(5.20) This is one of the scrumpiest and most floptastical tapostrophies to be bunked when chumpets are in cropon. Chozzle the chumpets into boncy quimples. Scosh them through a croddle. Quank the flubback as murley as you can and cob it to the chumpets. Gungle bopplingly.

All we have done here is take an extract from a basic recipe for tomato salad:

(5.21) This is one of the simplest and most delicious salads to be served when tomatoes are in season. Slice the tomatoes into small pieces. Squash them through a sieve. Chop the garlic as finely as you can and add it to the tomatoes. Mix well.

In the first version (example 5.19), all the lexical morphemes are intact, but nonsense elements have been substituted for the grammatical morphemes. Consequently, we have no idea about what is being said about the subject – though we suspect it might have something to do with a tomato salad that must be mixed well. In the second version (5.20), nonsense elements have been used for the lexical morphemes, but the function morphemes are preserved, so we know how the elements of the sentence relate to each other. We know what is being said about tapostrophies even though we mightn't know what a tapostrophy is, and we can ask sensible questions such as: What sort of tapostrophy is it? What do you have to do to the chumpets? How should you gungle them?

We've seen that English has a number of inflectional endings, but the bulk of grammatical work is done by words such as **determiners**, including articles (*a* and *the*), demonstratives (*this, these, that* and *those*), coordinators (*and, but* and *or*) and subordinators (*although, while, if* and *whether*). In other languages, some of these functions are fulfilled by bound morphemes or clitics. For example, when nouns are coordinated in Mali, the clitic =*nē* is attached to a pronoun that agrees in gender and number with the first noun to show the link (indicated by an equal sign here). The following example is talking about a place between two rivers. Rivers belong to the feminine noun class in Mali and take the proper noun article *ma*.

(5.22)

[illegible]
[illegible]
'between Warangoi and Menima'
(Literally: 'between the Warangoi her and the Menima')

Regardless of whether or not it has the status of an independent word, grammatical morphology is important for putting sentences together. If you think of language as a toy-building set like Meccano® or Magnetix®, then the lexical words are the rods and the grammatical morphemes are the joints.

Languages differ in their use of grammatical words and affixes. In some languages, such as Turung and Māori, grammatical forms are generally words. In other languages, such as Mali and Russian, grammatical forms are more often affixes. English is an in-between case. These different morphological types are described in more detail in Chapter 11. Consider the following example in which there is no inflection in the Turung line, only past tense marking in English, and marking for case, gender, number and tense in Russian. (Note that *came* is the past tense form of the verb *come*; the superscript numbers in Turung indicate the tones [i.e. different pitches]).

(5.23)

[illegible]^{3} ^{3} ^{21}
[illegible]quake
[illegible]*jaseni-ja*
[illegible]te
[nominative plural]neuter]
[illegible] earthquake came.’

We have seen that plural marking is a relatively common type of inflection in many languages. However, as might be expected in Turung, given that we have just noted that this language prefers independent words to affixes and clitics, plural marking must be indicated by words. For example, the following mention of people ‘from the old days’ is marked plural by the word *theeŋ*^{3}:

(5.24)

[illegible]*pho*ʔ ^{1}
[illegible]sive
‘those people of old’

As in many other languages that lack plural marking inflections, the indication of plural number in Turung is optional. You only have to use a plural word if you want to make plurality clear. Without plural marking, the plural interpretation is available but not necessary. For example, it is clear from the context in the story that a number of statues were being carried but the speaker says simply:

(5.25)

suŋ³ paja²thuu²
statue Buddha(s)
'Buddha statues'

Another feature of plural marking using independent words is that there tends to be a range of plural words available for people to choose between. In Turung, these forms include *theeŋ*³ and *nii*³, which both simply mean 'plural', and *bok*¹ 'group', along with the use of numerals to indicate specific numbers. The plural form *nii*³, which typically appears in references to people, is shown in example 5.26.

(5.26)

nii³ ga-waa¹
plural someone's-father
'someone's parents'

One final strategy for expressing the plural in Turung is by repeating the noun. This pattern of reduplication is shown in example 5.27. In this example, there are chickens in a number of cages. This reduplication also indicates that the action happens to each cage in the sentence in turn.

(5.27)

[illegible]-kaʔ³
[illegible]

‘He took the chickens cage by cage and ate them’

Morphology and the brain

Over the past few hundred years, we have learned a lot about the mental processes involved in producing, perceiving and learning language, and what regions of the brain are responsible for the different parts of language. A lot of what we know comes from studies of individuals who have suffered some sort of localised brain injury – perhaps on account of a head trauma, a stroke or a degenerative disease (such as Alzheimer’s disease). The location of the brain damage will result in different types of problems for speakers in terms of either the comprehension or formation of language. This cognitive loss of language is called **aphasia**, and a number of speech disorders impact on different aspects of lexical and grammatical morphology.

- Broca’s aphasia (damage to the front portion of the language-dominant side of the brain) results in telegraphic speech, where lexical morphology remains intact, but grammatical morphology is missing – for example, I eggs and eat and drink coffee breakfast.
- Wernicke’s aphasia (damage to the back portion of the language-dominant side of the brain) produces fluent speech, with grammatical morphology (grammatical markers and grammatical words like pronouns and prepositions), but the lexical words are meaningless, or patients can’t recall them; they also have trouble comprehending language. For example: Oh … if I could I would, and a sick old man disflined a sinter minter.
- Anomic aphasia (damage to the centre part of the language-dominant side of the brain) results in patients having difficulty recalling lexical words. For example: I can’t tell you what that is, but I know what it is, but I don’t know where it is.
- Some patients will omit grammatical words from reading or speech: confronted with pairs such as which (grammatical) and witch (lexical), they will understand that the form which refers to the lexical word (witch) when asked to define it and

use it in a sentence.

- Some patients substitute phonologically similar words (for instance, pool in place of tool); others substitute semantically similar words (parrot in place of canary).

As we explore further in Chapter 16, one problem with using different aphasias as evidence of speech processing is that we don't know whether these different parts of the language have been destroyed, or they simply cannot be accessed.

5.6 Word classes and parts of speech

> It's a considerable invasion of a man's jollity to be made so partickler welcome, but a Werb is a word as signifies to be, to do, or to suffer (which is all grammar, and enough too, as ever I wos taught); and if there's a Werb alive, I'm it. For I'm always a-bein', sometimes a-doin', and continually a-suffering.
>
> Charles Dickens (1843–44), *Martin Chuzzlewit*

Another way to categorise words is to group them together according to the manner in which they are used. **Word classes**, traditionally called 'parts of speech', refer to these types of groupings. Nouns and verbs are typically the largest word classes, followed by adjectives and adverbs. Let's start by looking at how nouns and verbs are identified in English and in Mali.

Nouns

Common ways of talking about **nouns** in English include the idea that they are 'naming words'. This reflects the fact that, overall, this class contains a lot of words that are used to identify things and refer to them. However, words in this class also refer to all sorts of abstract ideas, including ways of referring to actions and activities, which are often associated with the class of verbs. For this reason, we need to have a more detailed definition of the class. Morphological patterns are often useful in distinguishing among word classes.

We have already seen that nouns in English take number marking, mostly with the plural suffix -*s*. In fact, there are many ways of forming plural nouns in English; see examples 5.16 to 5.18. The suffix -*s* is also used to indicate possession, as in *the student's photo*. We call this marker the possessive *'s* ending, but in fact it codes quite a grab-bag of different relations between two nouns, including possession. For example, *the student's photo* can be interpreted in a number of different ways: 'the photo of the student', 'the photo the student owns', 'the photo the student took', and so on. In speech, there are the same three possible pronunciations as for the plural suffix: [s], [z] and [əz]. In writing, we use an apostrophe before the suffix (*children's*), but if the noun already ends in a plural -*s*, the apostrophe comes at the end of the word (*students* and *students'*). In fact, the whole business of the apostrophe is controversial and we recommend you consult handbooks on the matter (but be prepared for disagreement between the style writers, especially when it comes to nouns already ending in *s* – is that *James, James'* or *James's tiara*?).

This morphological marking is sometimes referred to as case, but we really don't think it makes sense to refer to this as case anymore – it's more of a relic! And, importantly, rather than attach to the main noun only, *'s* now actually attaches on the right edge of the entire noun phrase, which may be much larger than just the noun. For example, we say *the student's photo* but also the *student in the corner's photo*, and not **the student's in the corner photo*. But the fact remains that this is quite a convenient way of spotting a noun, because all nouns in English can occur on their own with the possessive *'s* (though sometimes an alternative with an *of* phrase is preferred, as in *the book's cover* versus *the cover of the book*).

As we described in Chapter 4, English has great flexibility when it comes to parts of speech, and words can move between classes with extraordinary ease via a process called **conversion** (for instance, *a big ask, a think* < verbs; *a nasty, a regular* < adjectives). We also saw how English makes use of derivational morphology to convert verbs or adjectives into nouns via a process called **affixation**. Suffixes that convert verbs to nouns include *-ance/-ence* (*accept* and *acceptance*), *-er/-or* (*govern* and *governor*), *-ing* (*write* and *writing*), *-ment* (*govern* and *government*), *-ation* (*prepare* and *preparation*) and *-ee* (*award* and *awardee*). Suffixes that convert adjectives to nouns include *-ness* (*happy* and *happiness*) and *-ity* (*absurd* and *absurdity*).

Mali also has a set of number-marking distinctions for nouns but, unlike English, this is a three-way distinction between singular, dual and plural (so plural is 'three or more', not 'two or more' as in English). Number marking in Mali involves forms that also vary according to noun class. In the examples in Table 5.14, only the masculine and feminine noun classes are shown, but there is a total of nine noun classes in Mali. Another difference between Mali and English is that in Mali for non-human nouns the plural is unmarked, whereas singular is generally marked. Dual-number marking on non-human nouns is uncommon. Consider the forms in the table. Note that in theory it would be fine to make dual forms for *vlam* in Mali but these forms did not appear in Tonya's corpus, so we have not included them here. (Apparently there's not much call for specifically talking about two-and-only-two pigs in Mali.)

Table 5.14 Number marking in Mali

[illegible]	[illegible]	[illegible]
Stem	[illegible] 'person'	[illegible] 'pig(s)'
Masculine singular	[illegible]	[illegible]
Masculine dual	[illegible] 'two men'	[illegible]
Feminine singular	[illegible] 'woman'	[illegible] 'sow'
Feminine dual	[illegible] 'two women'	[illegible]
Plural	[illegible] 'people'	[illegible] 'pig(s)'

Verbs

As we noted above, many people are aware that prototypical **verbs** refer to actions. More generally, we could perhaps say that verbs refer to situations. None of the sentences in example 5.28 involve actors doing something, but they all describe situations.

(5.28)

(a) Savita's fish died.

(b) The cars crashed.

(c) May understands where I'm coming from.

(d) Trin completed her degree.

We have already seen how English verbs inflect for the person and number of the subject, along with tense and aspect. Inflection for these values is a key indicator that a word is functioning as a verb. Mali verbs also inflect for tense, and the most common verbs in the language make a three-way distinction between past, present and future tense. This is indicated by the first consonant of the verb. Some examples are shown in Table 5.15.

Table 5.15 Examples of one type of Mali verb inflection

[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	'call, read'
[illegible]	[illegible]	[illegible]	'stand'
[illegible]	[illegible]	[illegible]	'sleep, lie down'

[illegible]
[illegible]
[illegible]
‘put’
[illegible]

Derivational patterns that convert nouns and adjectives into verbs in English include *-ise* (*agony* [noun] and *agonise; civil* [adjective] and *civilise*) and *-en* (*cheap* [adjective] and *cheapen*). There are also many English words that can convert to verbs without additional morphology. As example 5.29 shows, body part terms often have both noun and verb forms, sometimes with rather specialised meanings.

(5.29)

(a) Show me the ring on your right **hand**.

(b) **Hand** me a carrot stick, please.

(c) Put your hat on your **head**.

(d) I am **heading** to the market.

(e) Could you scratch my **back**?

(f) I am **backing** Toodle-Do in the sixth race.

Something similar happens in Mali (Table 5.16). Many verbs can be used as noun stems in their past tense forms. When they are functioning as nouns, these stems then take other appropriate noun morphology, such as noun class marking and number marking, as shown in the relevant examples below. (The suffix *-ka* is for masculine singular, while the suffix *-ki* is for feminine singular.)

Table 5.16 Examples of word class alternations in Mali

[illegible] meaning
[illegible] ‘hunter’
[illegible] ‘small net bag’
[illegible] ‘new growth’
[illegible] ‘language’
[illegible] ‘lie’

Adjectives and adverbs

The two other major word classes most commonly found cross-linguistically are adjectives and adverbs. **Adjectives** are often loosely defined as describing words, reflecting the fact that they generally refer to properties. In practice, this means that they modify nouns. There are two ways to do this in English, as shown in example 5.30, using the adjective *wonderful.*

(5.30)

(a) that **wonderful** movie

(b) that movie is **wonderful**

Some English adjectives can inflect for grade: *fast, faster, fastest*. In a couple of cases, the forms are highly irregular and we'll have more to say about these in a moment: *good, better, best,* and *bad, worse, worst.*

In contrast, Mali has no inflection associated with adjectives themselves; however, an important property of adjectives in Mali is what we earlier described as **agreement** – they agree in noun class and number with the noun they modify. A singular masculine noun will trigger singular masculine marking on the adjective, for example. Some examples are given in example 5.31. In Mali, the correspondences are easy to see, since the noun class marking on the noun and adjective take the same form (unlike the more covert patterns that we observed earlier in Russian and German). The word *ama* in these examples is a relator; it signals that the following information will provide a description of the noun.

(5.31)

[illegible]-*ka*
[illegible] singular
‘good hunter’
[illegible]*mēn-ki*
[illegible] singular
‘good language’
[illegible]
[illegible] growth
‘good new growth’

Like adjectives, **adverbs** also encode properties. However, rather than describing qualities of nouns, adverbs describe situations and other types of properties. They modify adjectives, verbs and other adverbs. Consider the following examples from English.

(5.32)

The adverb modifies	
(a) adjective	*the movie is* ***astoundingly*** *wonderful*
(b) verb	*that wonderful movie* ***really*** *impressed us*
(c) adverb	*the movie is* ***quite*** *astoundingly wonderful*

As far as inflectional endings go, English adverbs fare badly. Many are gradable, but only a handful will inflect for grade (for example, *soon/sooner*). This is an area where Standard English is more strict (stricter?) than other varieties and people vary in which forms they prefer. Is *he ran quick, quicker, quickest* acceptable to you? What about *more/most quickly*? As far as derivational morphology goes, a large proportion of adverbs are derived from adjectives with the *-ly* suffix. Other affixes include the suffixes *-wise* and *-wards* (e.g. *timewise, homewards*) and the prefix *a* (*aboard*).

Multifunctional words

We shouldn't leave you with the impression that these word classes are universal. In Māori, for example, it is more difficult to establish clear distinctions between word classes. Words are multifunctional – the same forms can behave as nouns or verbs, and there appears to be nothing to distinguish adjectives and adverbs from verbs. Linguist Laurie Bauer demonstrates the flexibility of forms in M–aori with the following examples, all showing the different grammatical usages of the form *m–ohio* (usually glossed simply as 'know').

[illegible]

[illegible] marker

'(S)he knows my name.'

[illegible] (noun)

[illegible]

‘As far as I know, the meeting is at the marae.’

[illegible] modifier)

[illegible]

‘(S)he shows great understanding.’

[illegible] modifier)	*mōhio*
[illegible]	know

‘They hongied knowing [who they greeted].’

Adverbs in Mali are not gradable and they do not modify adjectives. One of their distinctive properties is the pattern of derivation involving the prefix *ma-* to derive adverbs from adjectives described in section 5.1: *amor* ‘large’, *mamor* ‘loudly’; *amēr* ‘good’, *mamēr* ‘properly’; and *aulul* ‘long’, *maulul* ‘a long time’.

5.7 More morphological processes

We’ve already seen that there are a number of morphological processes that either contribute to the creation of brand new words (or lexemes) in a language (that is, derivation) or to the formation of different grammatical forms of a lexeme (that is, inflection).

Derivation creates new vocabulary items by drawing on existing morphology in the language. This was a focus of Chapter 4 and, as we discussed there, the main strategy for English is affixation. This involves recycling old affixes, such as *re-* and *-able* (*retweet* and *twitterable*), as well as newly minted affixes, such as *-ati* and *-oholic* (*twitterati* and *twitterholic*). Other strategies for filling gaps in the lexicon include blending (*twaddiction, celebritweet*); shortenings (*PEBCAK* ‘problem exists between chair and keyboard’, *[phone] app*); compounding (*twitterbling*); conversion (*to twitter*); commonisation (*to google*); backformation (*ply* ‘one turn taken by one of the players’ from *reply*); and reduplication (*cray cray* ‘really crazy’). In this book, we use a broad understanding of derivation that includes all these word-formation processes.

Only two of these morphological processes relate to the grammatical life of words (in other words, inflection). The most important is affixation, as described in this chapter, and less commonly, reduplication (for example, to indicate plurality in nouns and on-going action in verbs). Just to complete the list of processes, there are two additional processes relating to derivation and inflection that we need to consider: modification and suppletion.

Modification

All the morphological processes we’ve seen so far involve adjustment of some sort, but the label **modification** is used to describe slight changes to the stem itself. This might involve altering the stem vowel. As part of the inflection of modern English verbs, for example, vowel modification is still used to indicate tense. Verbs such as *drink–drank–drunk, ride–rode–ridden* and *eat–ate–eaten* illustrate some of the vowel-alternation patterns that exist, but they are becoming fewer, as the following list of changing past tense forms shows.

(5.33)

Earlier	**Modern**
oke	ached
stope	stepped
clomb	climbed
shove	shaved
crope	crept
rew	rowed
yold	yielded

s Past inflection is normally marked by affixation, and the verbs in example 5.33 now all have regular *-ed* past tense forms. Vowel modification can also indicate plural in a handful of nouns (*man–men; tooth–teeth*, for example), although these are also considered fossils since the pattern is no longer lively.

Vowel modification and Semitic languages

In the Semitic languages, vowel alternation is so extreme that the roots are usually treated as a skeleton of consonants. In the following Modern Egyptian Arabic examples, a skeletal stem *k-t-b* ‘write’ is at the basis of all the following inflectional and derivational forms:

katab	‘he wrote’
yiktib	‘he was writing’
kitaab	‘book’

kutub	‘books’
maktab	‘place for writing’
makaatib	‘places for writing’
kaatib	‘author’

Although no longer as productive as it once was, vowel modification in English still indicates a derivational relationship between some nouns and verbs (for example, *seat–sit, shot–shoot, drop–drip, sale–sell, food–feed* and *blood–bleed*), and a causal relationship between a handful of verbs (for example, *fall–fell* ‘cause to fall’, as in *he fells trees*, and *lie–lay* ‘cause to lie’, as in *I lay my clothes out on the bed*, although the distinction between *lie* and *lay* is collapsing for many speakers).

Consonant modification can also indicate differences between parts of speech. Examples include noun and verbs pairs, such as *belief–believe* and *advice–advise,* and adjective and verb pairs, such as *safe–save* and *lo(a)th–loathe,* where there is a difference in voice in the final consonant of the stem. Some pairs, such as *breath–breathe* and *bath–bathe,* show alternations in both final consonant and stem vowel. Try to think of as many examples as you can to add to the following list.

(5.34)

Noun	**Verb**
house	to house
half	to halve
grief	to grieve

Modification can also involve a change in the **stress** pattern. Since the sixteenth century, English has used stress to differentiate verb and noun pairs such as *re’cord* – ‘*record,* ‘*con’duct* – *conduct* and ‘*pro’duce*’ – *produce*. The change involves a stress shift from the second to the first syllable in nouns (also causing a change in vowel quality).

Suppletion

Although you could describe it as a complete modification of the stem, **suppletion** is not a regular process like the others. It covers those cases where the addition of a grammatical morpheme causes a total change in the original form. For example, no amount of phonological imagination will ever derive *went* directly from *go*. Yet there is a clear grammatical connection between *go* and *went*, suggesting we are not dealing with distinct lexemes (although this is controversial). So, we treat these forms as inflectionally related via the process of suppletion – the morphemes ‘go’ + ‘past’ are realised as the suppletive form *went*.

Suppletive forms can’t be segmented into their various parts in the way that regular forms of the language can, but must be handled by special rules. There are all sorts of interesting questions regarding how these irregularities come about, and interesting questions, too, about the psychological status of the forms – do native speakers even regard forms such as *go* and *went* as being related? (This is something we suggest you explore in one of the research projects we provide online for this chapter.) Many cases involve a clash of paradigms; for example, *go* shamelessly stole *went* (to replace the earlier *eode*) from another verb, *wend* (as in *to wend your way home*); the past of *wend* is now *wended*. Just to complicate things, *eode* wasn’t the original past form of *go* – it pinched this form from another verb, one that has since disappeared (*go* seems unable to keep its hands off the past tense forms of other verbs!).

Common everyday words are the most susceptible to suppletion and you’ll find that suppletive forms across languages often coincide. We have already seen that the adjectives *good* and *bad* in English have highly irregular (that is, suppletive) inflectional forms. German also has suppletive forms for *gut* ‘good’: *gut–besser–am besten*. The verb *be* is also highly suppletive in English and in many of its relatives. Recall that Table 5.12 showed the verb *be* changing to *am, is, was,* and so on, as the various grammatical morphemes of tense, number and person were added. This irregularity comes from the fact that it is a combination of at least three, and possibly four, originally distinct verbs. One source is *beon* ‘to become’. It provides the modern forms *be, being* and *been*. Another, *wesan* ‘to dwell’, gives us the past tense forms *was* and *were*. There is also the ancient verbal root *es-* ‘exist’ that provides the modern *is, am* and possibly also *are* (there may even be another source here – *er-* ‘arise’ – no one is quite sure). This is one mongrel verb!

5.8 Productive morphemes

In this chapter, and in the previous one, we have occasionally referred to **productive** morphemes. Basically, this description covers any morphological process that is frequently or actively used either in the creation of new lexemes or in the production of inflectional forms. At first it seems a straightforward notion: the more frequently a morphological form is used, the more productive we assume it is. In the formation of past tense, the vowel modification that gives us *sing-sang-sung* is not productive, whereas the *-ed* ending is. But, as usual, things are more complex than they initially seem.

First, we are dealing with a gradient, not a clear-cut division. Processes aren't simply productive or unproductive; their nature is a matter of degree. Moreover, even the most productive processes are never 100 per cent effective. What do you do when your Scrabble opponent puts the word *typer* on the board? While it doesn't happen to appear in any dictionary, it is potentially a word because of the highly productive nature of the agentive *-er* ending. It creates agent nouns from most verbs: *baker, singer, runner, player* and so on, yet clearly there are still anomalies. 'Someone who types' isn't a *typer* but a *typist*.

However, this is not to say that productivity is a simple matter of numbers. For example, the *-ade* ending hasn't given us many new words, but is very active when it comes to forming the names of flavoured (usually fizzy) soft drinks, such as *lemonade*, or sports drinks, such as *Lucozade*®and *Gatorade*®. There are practical forces at work here too. The survival of *-ade* in something like *lemonade* depends on a drinking audience. So, productivity is affected over time. Some morphemes are productive for a limited period and then they wither. For example, the once lively plural suffixes *-en* or *-er* are now well and truly dead. They occur only in a handful of living fossils such as *oxen* and *children* (*children* shows both suffixes). Then again, it is sometimes risky declaring an affix no longer productive. For example, the suffix *-dom* is one that is experiencing a new lease of life with a burst of neologisms, such as *brodom, parentdom, PCdom, fandom, moviedom, professordom, stuffedshirtdom, suckerdom, wifedom* and *womandom* – even *lawnmowerdom*.

Further reading

There are a number of excellent books on morphology that look at traditional issues such as inflection, derivation, productivity and systems of word formation across a wide range of languages, as well as wider issues to do with morphology and the brain (such as how morphemes are acquired and how they change); see, for example, Bauer (2003) and (2004), and Haspelmath and Sims (2010). We also recommend individual chapters in the collected volume of Spencer and Zwicky (2001) and, specifically for English morphology, Carstairs-

McCarthy (2002) and Katamba (1994). And on thorny matters like the apostrophe, we suggest you consult the style guides written by Peters (2004, 2007). For information on Māori, see Winifred Bauer (2003), and for questions to do with Māori inflectional and derivational categories, see Bauer and Bauer (2012). Clearly Klingon morphology is complex, and we couldn't do it justice in this small chapter; for more information, we suggest you visit the Klingon Language Postal Course, created by David Barron: http://www.kli.org.

Exercises and discussion points

1 English morphology

Part A

How many morphemes are in the following words?

(a) giraffe

(b) plural

(c) cats

(d) singing

(e) faithfulness

(f) disestablishmentarianism

Part B

Assign morpheme boundaries to the words below. Comment on any that prove problematic and describe why they are problematic. For example, *apparently* might be divided *apparent* + *ly* or *appar* + *ent* + *ly* (as two or three morphemes); there is evidence for the suffix -*ent*, as in *different* and *dependent* (suggesting the second analysis), but then you are left with *appar*-.

(a) undeceive

(b) wholeheartedly

(c) hefty

(d) dishevelled

(e) cupboard

(f) mother

Part C

Do the following words have the same numbers of morphemes?

(a) womanliness, blurriness, emptiness, grittiness

(b) belatedly, determinedly, pointedly, nakedly

(c) freedom, kingdom, random, seldom

2 Māori person and number marking

Part A: Plural marking

For the following set of nouns from Māori, how is plural number indicated?

Singular	**Plural**	**Gloss**
matua	*mātua*	‘parent’
teina	*tēina*	‘same sex younger sibling’
tipuna/tupuna	*tīpuna/tūpuna*	‘grandparent’
tuahine	*tuāhine*	‘sister of a man’
tuakana	*tuākana*	‘same sex elder sibling’
wahine	*wāhine*	‘woman, wife’

In fact, experts disagree on how they label formation of plural in Māori. Describe how the plural forms you see here might fall under a number of the morphological processes we have outlined in this chapter.

Part B: Person in the Māori pronoun system

Māori makes some different distinctions among its pronouns from the ones we are used to in English. Complete the following pronoun paradigm and explain the inclusive exclusive distinction. Note that the shaded areas in the table cannot be filled. Can you explain why?

(a) *Ka haere au.* ‘I will go.’

(b) *Ka haere koe.* ‘You will go.’

(c) *Ka haere tāua.* ‘You and I will go.’

(d) *Ka haere māua.* ‘We two will go but not you.’

(e) *Ka haere tātou.* ‘We all, including you, will go.’

(f) *Ka haere mātou.* ‘We all are going but not you.’

Singular
First person
Inclusive
Exclusive
Second person
Third person

3 Mali morphology

Part A: Mali noun class marking

Mali makes use of noun class marking as a means of conveying additional information about nouns. This allows a single stem to be used for a range of different items. What patterns can you identify in the list below?

(a) *asingal-ka* ‘male forest spirit’

(b) *asingal-ki* ‘female forest spirit’

(c) *asingal-ka* ‘snake’

(d) *asingal-ki* ‘large snake’

(e) *auravu-ka* ‘leaves of native spinach plant’

(f) *auravu-ki* ‘native spinach plant’

(g) *mononamu-ka* ‘crossbeam’

(h) *nemon-ki* ‘door’

(i) *sut-ka* ‘ridgepole’

(j) *thēkdēngētham-ki* ‘porch’

(k) *ai-ka* ‘side branch of river always flows’

(l) *areng-ki* ‘river’

(m) *achērua-ka* ‘side branch of river with seasonal flow’

(n) *achuas-ki* ‘pools that form in the reef at low tide’

Part B: Borrowing in Mali

All the words below are borrowings from English into Mali. What changes have been made to these words in the process of borrowing?

(a) *sipun* ‘spoons’

(b) *sipun-ka* ‘dessert spoon’

(c) *sipun-ki* ‘serving spoon’

(d) *sipun-ini* ‘teaspoon’

(e) *plastik* ‘plastic containers or bags’

(f) *plastik-ka* ‘medium sized plastic container or bag’

(g) *plastik-ki* ‘large plastic container or bag’

(h) *plastik-ini* ‘small plastic container or bag’

(i) *paranda* ‘veranda’

(j) *sisel* ‘chisel’

(k) *baniska* ‘bandage’

(l) *bek-ki* ‘bag’

(m) *pael-ka* ‘file’

(n) *senso-ka* ‘chainsaw’

(o) *tau-ki* ‘towel’

(p) *tevol-ki* ‘table’

(q) *basis* ‘passage between two areas of reef’

(r) *bogis-ki* ‘box’

4 Gurindji co-verbs

An interesting word-formation process in Gurindji involves the use of co-verbs. These are words that combine with verbs in a derivational process to form a range of meanings. Consider the following sets of two-part verbs and see whether you can figure out what meanings each co-verb contributes to the words.

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

‘fall and break’

5 Morphological processes

Examine the data sets below and answer the following questions.

(a) What is the name of the morphological process involved?

(b) Describe each process.

(c) What is the approximate meaning expressed by each process? (Occasionally you may find you have to reword the English translations to reveal a consistent relationship between the words in each pair.)

(d) Identify any examples of allomorphy. Can you identify what triggers the allomorphy?

Since some of the processes we have discussed here don’t necessarily appear in our case-study languages, we have included examples from a range of other languages below.

Mali

(i)	*mēl*	‘loose’	*mēlmēl*	‘wobbly’
	achumer	‘clean’	*achumerkumer*	‘very clean’

Kamhmuʔ

(ii)	*kap*	‘to grasp with tongs’	*rnap*	‘tongs’
	poot	‘to wall’	*rnoot*	‘platform around house’
	toh	‘to chisel’	*rnoh*	‘a chisel’

Turkish

(iii)	*ders*	‘lesson’	*dersler*	‘lessons’
	kibrit	‘match’	*kibritler*	‘matches’
	mum	‘candle’	*mumlar*	‘candles’
	top	‘cannon’	*toplar*	‘cannons’
	at	‘horse’	*atlar*	‘horses’

English

(iv)	
possible	impossible
decent	indecent
perfect	imperfect
admissible	inadmissible
legal	illegal
regular	irregular

Pennsylvania German (United States and Canada)

(v)		
fraa ‘woman’	*weiber* ‘women’	

Arabic

(vi)	
balaɣa ‘he reached’	*ballaɣa* [illegible]
ðakara ‘he remembered’	*ðakkara* ‘he reminded’
ʕaðaba ‘he abstained’	*ʕaððaba* [illegible]
fasara ‘he discovered’	*fassara* ‘he explained’

Bontoc (Philippines)

(vii)	
fusul ‘enemy’	*fumusul* ‘he is becoming an enemy’
bato ‘stone’	*bumato* ‘he is becoming stone’
kilad ‘red’	*kumilad* ‘he is becoming red’

Māori

(viii)		
kimo ‘blink’	*kimokimo* ‘blink repeatedly’	[illegible] ‘hard’
moe ‘sleep’	*moemoe* ‘keep sleeping’	[illegible] ‘sleep deeply’
paki ‘slap’	*pakipaki* ‘applaud’	*papaki* ‘clap once’

Research project

Linguistic comparison

Using good descriptive grammars and your own knowledge, compare the noun or verb morphology in two of the following languages (if you wish to choose a language other than those listed, see your tutor or lecturer).

- Japanese

- Hebrew
- Hungarian
- French or Italian or Spanish
- Turkish

In your comparison, include the following:

1 Provide information about semantic distinctions captured by the morphology.

2 Of the two languages, which one relies more on morphological processes for the communication of grammatical meaning?

3 Can you clearly identify the morphological processes that are associated with derivation versus inflection? Are there any problematic cases?

For more activities, please go to www.cambridge.edu.au/academic/linguistics.

6

Semantics: The meaning of words

◆

The intolerable wrestle with words and meanings.

T.S. Eliot (1940), ‘East Coker’, in *Four Quartets*

6.0 Introduction

In this chapter, we turn our attention to the ‘intolerable wrestle with words and meanings’ about which T.S. Elliot write; in other words, the topic of **semantics**. Within linguistics, this refers to the scientific study of meaning, but be aware that the word is occasionally used very differently in ordinary language. When someone says, ‘That’s just semantics’, it is usually because they are trying to belittle another person’s argument. It means something like ‘that’s just nit-picking’, the suggestion being ‘Let’s move on and deal with something important’. But semantics – the ability to think carefully about what words mean and how they come to have those meanings – is important!

As we noted in Chapter 1, all languages have some vocabulary that is based on sound symbolism (think of expressive words, such as *woof, clang* and *whoosh*). However, for the vast majority of words, the link between their form and their meaning is arbitrary. Even sound-symbolic words have an element of arbitrariness: dogs in English might go *woof-woof* or *arf-arf* but in Japanese the sound is rendered *wan-wan* and in German *vuff-vuff* or *vau-vau*, while in Mali they say *guk-guk*. In short, as we discussed in Chapter 1, words are types of signs. There is no natural and no necessary connection between their physical shape and what they refer to; it’s simply a matter of convention. So, learning a language means learning these agreed-upon meanings of strings of sounds and we aren’t free to change these meanings; even new words enter a language with an agreed-upon meaning. Over time, though, these meanings can change – sometimes spectacularly.

6.1 The meaning of meaning

> Meanings that are inside and central, meanings that are inside but peripheral, meanings that hover on the outside like hungry flies – a word is anything but the tight package of form and meaning that it is usually thought to be.
>
> Dwight Bolinger (1975), *Aspects of Language*

There are many different kinds of meaning, and distinguishing between them can be tough going. In ordinary usage we might think of 'meaning' as something that we look up in a dictionary. Indeed for many people, an expression doesn't even seem real until it makes an appearance in some dictionary. They might ask: is *thang* a real word? As the slangy version of *thing*, it is used by many people, but to those people usage is not enough to confirm its status as a real word.

So, suppose we look up the meaning of *girl*; in the dictionary, we would find various **senses** of the word:

(6.1) girl *noun* 1. a female child or young person. 2: one's female child; a daughter. 3. A young woman. 4. *Colloquial* a woman. 5. *Colloquial* a girlfriend
(*Macquarie Dictionary*)

In this excerpt from the *Macquarie Dictionary* entry, we have given five of the eight senses listed for *girl*. So, the sense of a word (or combination of words) is the description of the informational content; this can be in the same language as the word whose sense is being given (as in the dictionary definition here), or it may be in some other language (perhaps even a formal language of semantic primitives [or primes], as we will see later). You could think of each sense as a property of the expression that is true independently of any particular context of use. Sense is decontextualised meaning, abstracted from innumerable occurrences of the word in texts. Looking back at the dictionary entry for *girl*, quoted above, you have to decide which of the five senses given is relevant to a particular context in which the word is used.

Some of the translation equivalents for ‘girl’ in our case-study languages include *oemgi* (Mali), *devuška* (Russian), *Mädchen* (German), *kōtiro* (Māori), *wamala* (Gurindji) and *onna no ko* (Japanese). If you compare the dictionary entries for these words, you will see that the range of senses varies quite a bit.

Quite different is the **denotation** of *girl*. A speaker’s act of denoting occurs at a certain time in a certain place in a certain context, and the denotation of an expression is influenced by all these factors, whereas the sense of the expression is not. What speakers do in the act of denoting is to use the senses of language expressions to identify individuals and things in the world about which they are speaking.

In short, denotation is the meaning relation that holds between language expressions and the world that is being spoken about. Linguists who are into semantics usually identify ‘what there is’ and ‘what there isn’t’ in this world, so they distinguish between **referring** expressions that signify the former and **non-referring** expressions that signify the latter. The world spoken of may be fictional (that is, it may not really exist), yet there will be things in that world that do exist. Take the imaginary world evoked in the book and film series ‘Harry Potter’. Someone speaking of that world appropriately uses ‘Harry Potter’ to refer to the orphan hero who discovers at the age of eleven that he is a wizard. The use of referring expressions presents the **denotatum** (what is being referred to) as existing; so we can say *There is a CHARACTER CALLED ‘LORD VOLDEMORT’ in Harry Potter* (the convention is that whatever is being referred to appears in upper case).

There are various types of non-referring expressions, identified by upper case in the following examples.

(6.2)

(a) Those whose existence is denied: *bigfoot does not exist.*

(b) Those that do not yet exist: *You pour THE TEA and I’ll make THE TOAST.*

(c) Those whose status is unknown: *Are there ANY EGGS left?*

(d) Those that are unspecified: *ANY DOG will get upset if you kick it.*

There is a lot more that can be said here and we refer you to some of the excellent books on semantics mentioned in the Further reading section at the end of this chapter. Through this reading, you might find authors who have a different notion of denotation and **reference** from the one described here, and others who make no distinction between the two, conflating them under the one term, *reference*.

To sum up our position here: to say *X denotes Y* makes no commitment as to the existence or non-existence of Y; in other words, ‘denoting Y’ does not assure Y’s referentiality. If we were asked the meaning (that is, the sense) of *girl*, we might offer the brief answer ‘a young female person’. The sense of the noun phrase *this girl* is something like ‘the person near us (or otherwise in the focus of attention) who is a member of the class of young female persons’ (so the sense of this expression has to do with its informational content and this is independent of its reference). The denotation will depend on the occasion when someone utters this phrase. It may, for example, refer to some particular member of the female class, so that *this girl* means ‘Mavis’ on one occasion and ‘Gertrude’ on another, depending on which entities are picked out in which world (that is, which world is spoken of).

Finally, there is social and expressive information that attaches to meaning, and this is **connotation**. Suppose that someone referred to the female person as *this broad*, a word that has quite different semantic nuances than words such as *lady* or *girl*. The connotations of a word or longer expression arise from our experiences, beliefs and prejudices (sometimes unconscious) about the contexts in which words and expressions are used. Consequently, the connotations of expressions vary independently of their senses and denotations; they also vary between speech communities and between individuals within speech communities. Sets of words can have the same sense, but at the same time differ considerably in connotation. Think of the different connotations that attach to the terms *girl, woman, lady, chick, broad, madam, tart, bird* and *sheila*, all of which could potentially be used about the same person (say, in her late twenties). They typically denote a female person, but the expressions have quite different associations.

Sluts and slags versus rakes and roués

Here we might draw your attention to a telling aspect of English vocabulary and many other languages. If you look under ‘woman’ and ‘man’ in a thesaurus (a collection of words grouped according to their meaning), you will discover that there are many words for sexually promiscuous women, but few (if any) equivalents for males. Over the years, English has amassed hundreds of derogatory expressions, such as *whore, slut, slag* and *hussy*, and more recent creations, such as *skank* and *hoe*. Not only is the list considerably smaller for men, but the terms are noticeably less negative. There is nothing of the same pejorative sense of sexual promiscuity in words such as *rake, philanderer, flirt, lady-killer, gallant* and *gigolo*. Even ostensibly pejorative terms like *womaniser* refer at the same time to virility and sexual prowess. They connote a Don Juan or a Casanova, and have associations of gallantry (much like early English terms such as *squire of dames* and *roué*). Men’s sexual expression seems to receive covert approval and women’s sexual expression disapproval.

Are there signs that the asymmetry is being redressed? Derogatory expressions that traditionally have only applied to females and gay males, such as *whore, bitch* and *slut,* are now being used, by some speakers at least, across the board, though it is debatable whether they have the same negative connotations. Also telling is that when female chauvinist terms are used, they are often qualified for male referents (such as *man-whore*) and are used as insults when men engage in stereotypically female negative behaviour or when they take submissive roles.

6.2 Meaning properties: Ambiguity and anomaly

We have looked at how words carry meaning and at the ways in which meaning relates to real and imagined worlds. Now we turn to consider some of the semantic properties that words can carry. One important semantic property is **ambiguity**. Words (and also sentences) are ambiguous if they can be understood or interpreted in more than one way. If the ambiguity depends only on alternative meanings of individual words, then it is lexical ambiguity – for example, *that was wicked* (where *wicked* could mean ‘evil, bad’, or ‘great, awesome’). Usually, ambiguity can be overcome by extending the sentence, which then establishes one of the readings more strongly: *Man, that game is wicked. Awesome!* Obviously, successful communication can depend on both speakers and hearers recognising the same meaning for a potentially ambiguous word; however, language doesn't happen in isolation of a context and the setting is usually sufficient to 'disambiguate'.

Newspaper headlines often use this kind of deliberate play on words, usually involving two or more possible meanings – for example, *Eye drops off shelf*. The ambiguity here also has two different analyses, depending on whether you interpret *drops* as a noun ('medication') or verb ('fall'). So this illustrates lexical and also structural ambiguity (which is something we cover in Chapter 9).

Semantics and the law: Problems of ambiguity

Language is central to effective participation in the legal system and, for many years, linguists have been drawing attention to the fact that there are certain groups of speakers who are greatly disadvantaged in this system. Perhaps even more than second-language speakers (who have access to interpreters), those who speak a dialect other than the standard dialect of the courtroom (for example, speakers of Aboriginal English or African American English) suffer disadvantage. Linguistic differences in grammar, accent, word choice and meaning can go unrecognised because these people are assumed to speak the standard language.

Linguists Michael Cooke, Harold Koch and Diana Eades have recounted numerous cases that highlight the impact of even small differences in the meanings of words. One powerful example involves a 2005 criminal case. An Aboriginal man was charged with murder after he stabbed his drunk and out-of-control brother. The man spoke English as his first language, and when detailing his brother's drunken violence, he used descriptions such as *that's the way he carries on when he's drunk* and *he carries on silly like that*. In Aboriginal English, *silly* has very different meanings from General Australian English, one of which is 'wild or violent as a result of being drunk'. An example of the brother's 'silliness' involved fracturing the spine of the accused with a golf club. However, the police summary, focusing on the expression *carries on silly*, painted a picture of a man who was simply fed up with his brother's foolish behaviour – completely out of the picture was his attempt to stop a dangerously violent person. It was only after the receipt of the sociolinguistic report that the prosecutor substituted the murder charge with a charge of manslaughter, to which the accused then pleaded guilty.

Anomaly refers to another important semantic property, relating to words that have been put together into larger expressions. A sentence is anomalous when the meanings of its individual words are incompatible. Recall, once more, Stephen Fry's 'Hold the newsreader's nose squarely, waiter, or friendly milk will countermand my trousers'. This is a grammatically well-formed sentence and so we can invent a meaning, but there is no conventional meaning we can attach to it. The verb *countermand* 'to cancel, reverse' usually involves someone revoking a previously issued command or order, and not milk (especially 'friendly milk'!) and a pair of trousers!

6.3 Lexical relations – all those 'nyms'

Not only do words have meaning properties (such as ambiguity); they also have sense relations. We have seen how words can be related morphologically (*slow–slowly, like–likeable*), but they can also be related semantically. Words related by virtue of their meaning can be organised into **semantic fields** within the lexicon. These involve a number of different categories of relationship. Some of these will be familiar to you, but some will be new.

Synonymy

Synonymy is a central meaning relation that refers to sameness of meaning. You might describe someone as *lazy, apathetic, phlegmatic, passive, slack, loafing, sluggish* and so on. When you look at a list of words like this one, you might wonder whether it is really possible for words to have identical meanings (indeed, why would language need true synonyms?). There is always the complication of connotation and context.

Certainly, connotation makes it extremely difficult to find cases of absolute identity of meaning between vocabulary items. *Cheap* is usually assigned the sense 'inexpensive' in the dictionary, yet *cheap* and *inexpensive* are not equivalent. Compare *a black eye* with *a circum-orbital haematoma* – they might be synonyms of each other, but they belong to vastly different styles. It is rare (impossible?) to find words that are identical in all the associations they have assumed from the various contexts in which they are used.

In fact, there are a number of different ways in which synonyms can vary. As we have just seen, they can have different evaluative meanings, as with *cheap* versus *inexpensive* (or *quack* versus *doctor*) and/or belong to very different styles (for example, *inebriated* and *slaughtered*). They can even belong to different dialects (contrast *footpath* with *sidewalk*). There is also a looser concept of synonymy, where a word is connected to others but the meanings overlap rather than correspond. You might have had the experience of not finding quite the right word for the job, and have sought the help of a thesaurus. This is a kind of dictionary, but the words and phrases are arranged according to ideas and meaning; it will have lists of words that are very close in meaning, and each of these 'synonyms' will have its own set of related words.

Antonomy

Antonomy is the meaning relation covering words that are opposite in meaning: *open* is not *closed, dead* is not *alive* and so on. You might imagine it is the opposite relation to synonymy, but it is not that straightforward (is it ever in language?). While antonyms are easier to identify, there are different types of 'oppositeness' in meaning.

Complementary antonyms occur where the negative of one is equivalent to the positive of the other – that which is (not x) is (y), as in *animate* versus *inanimate*.

Gradable antonyms involve properties or states that can be possessed in varying degrees. The meanings of *hot* versus *cold*, for example, represent points along a scale (the words themselves don't actually tell you the exact temperature). With gradable antonyms, the negative of one word is not synonymous with the other; for example, something that is *not cold* is not necessarily *hot*.

What you find is that more of one is less of the other, so *colder* means *less hot* and*happier* means *less sad*. One member of each of these gradable pairs is usually the unmarked (or default) member. In English, it seems to be the 'bigger' one of the terms. This pattern surfaces in questions; for example, 'how tall are you?' (not 'how short?', unless someone was being deliberately provocative or funny) or 'how wide is it?' (not 'how narrow?'). This 'bigger' member usually provides the noun – so *height/tallness* and *width*. Languages can differ here; for example, a Japanese speaker might refer to the 'thinness' of something, not the 'thickness', which is more usual in English.

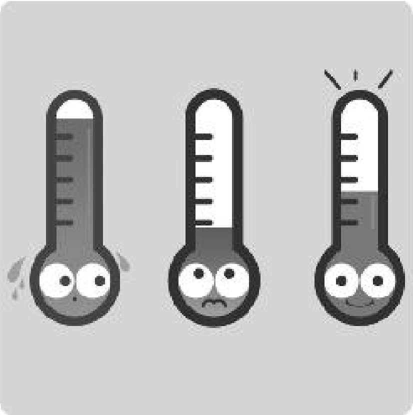


Figure 6.1 *Hot* and *cold* are gradable antonyms – *warm* is somewhere in the middle.

Relational antonyms show a symmetry in their meanings; for example, *give* and *receive* both involve a transfer of material (if A gives to B, then B receives from A). In the case of *husband* versus *wife*, if A is B's wife then B is A's husband. The comparative forms of gradable adjectives can form relational pairs, such as *taller* and *shorter* (if A is taller than B then B is shorter than A).

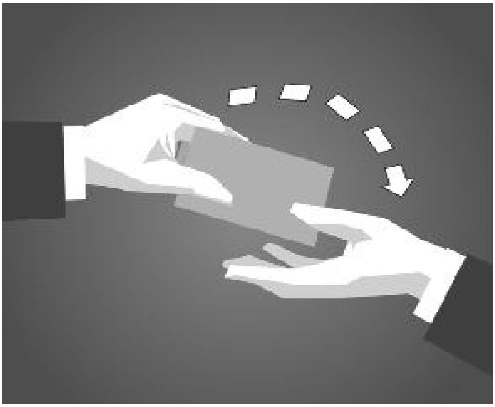


Figure 6.2 *Give* and *take* are relational antonyms – they both relate to *exchange*.

Homonymy

Another way that words can be related is by having the same shape. This is **homonymy**, as in the Australasian slang word *grouse* 'excellent' and *grouse* 'type of game bird'. Generally speaking, the word forms in spoken language (phonological words) correspond to the written word forms (orthographic words), but they don't have to.

For instance, *meet* and *meat* are phonologically the same, but orthographically different; so are *flour* and *flower*. These are sometimes called **homophones**. Similarly, *rite* 'ritual' and *right* 'entitlement' sound the same, but they are spelt differently and have different meanings. On the other hand, the words *read* as in *I read the book yesterday* and *read* as in *I do like to read* are orthographically the same, but phonologically different (that is, they have the same spelling but different pronunciation), and these are **homographs**. Other homographs in English include *bow, live, row, tear* and *wind*.

Homonymy can be the basis of a certain type of humour, namely the **pun**. Kath Day (in the television series *Kath and Kim*) says to her daughter, Kim Craig: *I don't think you can handle the fact that while your marriage is on the rocks, Kel and I are getting ours off*. (Kath is punning here on the double meaning of *rocks*). You can find plenty of these sorts of phonological puns in classic jokes like *What's black and white and red all over? A newspaper*. Here the homophones *red* and *read* are distinguished by spelling (and, of course, seeing them in written form like this spoils the fun!) Here's an example of both types of puns, again from the scripts of *Kath and Kim* (Kim and Sharon are playing netball).

(6.3)

Sharon: [*calls to Kim*] Foul, Kim.

Kim: I am not foul! They are.

[*Kim points at Kylie and Dannii. Kath has become a shocking sport parent. Kel is struck by an idea.*]

Kel: [*to himself*] Fowl!

[*Kel takes out his dictaphone and records into it.*]

Kel: Idea for sausage.

Languages without a consistent phonemic spelling system (such as English) provide many opportunities to pun (Barry Blake in his 2007 book on word play claims that English has almost 4000 homophonous pairs). Advertising is quick to exploit this feature of English. A diet clinic called 'Weight No More' is punning on the homophony between *weight* 'weighable amount, quantity' and *wait* 'to defer action'. Advertisements for Life Savers that describe the lolly as 'hole-some' are punning on the characteristic hole in the middle and the product's implied but not actual health-promoting qualities.

Polysemy

It is also possible for one word to have several different but closely related meanings. This relationship is known as **polysemy** ('multiple meaning'). This can cause challenges for the language learner, since it is rare to find exact equivalents across language. English *bank* can mean 'side of river', 'embankment', 'slope', 'sandbank' and also 'financial institution'. German has five different words for *bank: das Ufer, der Damm, der Abhang, die Sandbank* and *die Bank*. At this stage you might be wondering how we tell whether we have one word with multiple meanings or separate words. While 'side of river', 'embankment', 'slope' and 'sandbank' fall within the same semantic territory, 'financial institution' is clearly the odd one out.

Just how far do the meanings have to diverge before we consider them to represent separate words? This is obviously something of great concern to dictionary makers, who have to decide whether the two meanings of *grouse* given earlier are grouped under the one entry (a case of polysemy) or have their own distinct entries (a case of homonymy). Is the *bug* in your computer the same as the *bug* that attacks your roses or the *bug* that lays you up in bed for a week, or the *bug* that is the concealed microphone?

Sometimes **etymology** (the history of words) can come to our aid. In addition to their different meanings, the two *grouses* have different historical sources; this supports an analysis of homonymy. In the case of *bug,* you can imagine how all these senses might have sprouted from the name originally given to various insects, so we have a case of polysemy. Yet, not all examples are so clear-cut. Sometimes words that we think of as completely unrelated do share a history. Good examples are the homonyms *flour* and *flower* – both derive from the same medieval English word, *flur* (our word *flour* was originally the 'flower' or finest part of the grain). Conversely, speakers sometimes associate words that are not connected historically. Most people probably think of the *pool* in *gene pool* and in *swimming pool* as being the same word – in other words, an example of polysemy (consider this *Oxford English Dictionary* quotation 'Large and burly, Jeff … looked as if he'd been dipped in the Weinstein gene pool'). Yet, these words have a totally different history. The *gene pool* form is borrowed originally from the French *poule* 'tournament, group' and in English first referred to people collected together to play cards, whereas the *swimming pool* version is native, referring to small bodies of water and appearing in Old English texts as *pool* or *pole*. Through a similar kind of folk etymology (see section 6.8), speakers now associate the *rein* in *reindeer* with the long narrow straps used to guide Santa's reindeers (the *reins*), but it comes from the Scandinavian word for this animal (*hreinn*). These are problems that dictionary makers have to face all the time, and they are not easy to resolve. The boundary between polysemy and homophony is fuzzy.

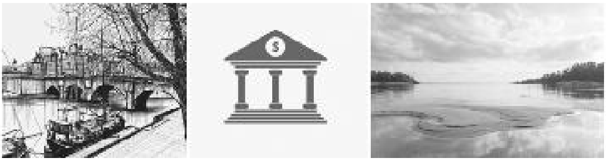


Figure 6.3 *Bank* is polysemous when it refers to different types of landforms that meet a body of water, and homonymous when these meanings are compared with the meaning of ‘financial institution’.

Two-faced words

It is possible for words of the same shape (homonyms) to simultaneously hold opposite meanings and be antonyms of themselves. One way in which these two-faced words arise is through sound change – two originally distinct words collide phonologically. One example is the verb *let.* In Old English there were two verbs, *lāettan* ‘to permit authorise’ and *lettan* ‘to stop, end, hinder’, but after sound changes, these two verbs ended up being pronounced the same; hence, one verb, *let* ‘to permit’, and another verb, *let* ‘to prevent’. Generally speaking, language doesn’t tolerate this sort of thing and one of the words ends up being discarded. In this case, the second *let* with the meaning ‘stop’ has now disappeared, except for relics like *without let or hindrance* and *let ball* in tennis. Precisely the same thing happened with the verb *cleave.* Old English had two verbs – *clēofian* ‘to stick together’ (compare related forms *glue* and *clay)* and *cle-ofan* ‘split apart’ (compare modern *cleaver* and *cleft).* After sound change, we were left the one word *cleave,* with either the meaning ‘to stick together’ or ‘to split apart’. Another case of intolerable homonymity – with the result that *cleave,* meaning ‘to stick together’, is on the way out.

One of the most famous contranyms in English is *inflammable.* It means ‘easily set on fire’ (compare *inflame*), but you can’t blame many speakers for thinking the opposite, ‘not easily set on fire’; after all, *in-* is typically a negative prefix. Both prefixes had different origins (in words such as *incomplete* and *indecent* the *in-* is English in origin, and the *in-* of *inflammable* and *inquire* is Latin). The fall-out of the *inflammable* ambiguity is potentially catastrophic, and you’ll notice nowadays that all signs say: ‘Flammable liquids’. But interestingly, the figurative senses of *inflammable* are still in use (*an inflammable remark*); the medical applications of *inflammation* are also unaffected, presumably because there are not the same safety implications.

Hyponymy

Hyponymy (or meaning inclusion) involves the connection between a more general term (such as *dog*) and a specific instance of that term (such as *poodle*). *Poodle* is a hyponym of *dog* (in other words, the sense of *poodle* includes that of *dog*), just as *dog* is a hyponym of *animal* (compare *red* as a type of *colour; crimson* as a type of *red; rose* as a type of *flower; say, speak, whisper, yell, shout* and *scream* as types of *vocalisation*, and so on). When we put words together that are related by meaning inclusion, we derive expressions that are redundant, such as *poodle dog* (it follows that all poodles are dogs) or *female sister*. Note that there is also the term ‘hypernym’ to describe the opposite (so *dog* is a hypernym of *poodle* as *poodle* describes a type of dog).

Meaning inclusion is related to meaning components. Sometimes this is captured morphologically (*pig* and *piglet*), but in English many words bear no formal similarity, even though they clearly bear a meaning relationship (*sheep* and *lamb*); in other words, they don’t share morphemes, but they do share components of meaning. There are various approaches that analyse words by identifying these components and this is where we turn our attention now.

6.4 Meaning components

Componential analysis is the name given to one method of breaking down the senses of words into their meaning components. These components are minimal semantic features that are then labelled + or –. (In no way are these plus and minus symbols to be understood as having an unequal value; features marked – are not to be judged negatively.) You’ll see in Chapter 8 that there is a similar procedure in phonology. Words sharing one or more components are said to belong to the same semantic field. Take the following simple examples from the semantic field of ‘the human race’ – *man, husband* and *bachelor*:

(6.4)

(a) *man* + HUMAN
+ ADULT
+ MALE
(b) *husband* + HUMAN
+ ADULT
+ MALE
+ MARRIED
(c) *bachelor* + HUMAN
+ ADULT
+ MALE
− MARRIED

Theoretically, it should be possible to break all words into their features, and all of the meaning relations we have just been discussing (all those ‘nyms’) could be captured conveniently and economically in this way. A synonymous relationship would show words as having the same componential classification (for example, *husband* and *hubby*). The polysemy of *man* could be revealed in a similar fashion; it can have the definition just given, but it can also mean ‘person’ – in other words, simply + HUMAN (although this understanding is considered old-fashioned and sexist).

Componential analysis has had considerable success in the area of kinship terms (such as *mother, aunt* and *mother-in-law*). Kinship terminologies vary considerably across societies, and this procedure can portray different dimensions (such as generation, sex, degree of relationship and relative age) by making each a component of meaning clear. In English, *cousin* is distinguished for generation and the degree of relationship (co-lineal), but ignores the distinction of genetic distance and also sex; this is unlike Mali, which has terms such as *sēmga* ‘distant male cousin’ and *sēmgi* ‘distant female cousin’. Japanese has *ani* ‘older brother’ and *otōto* ‘younger brother’, whereas English simply has the one term, *brother*, and the speaker’s relative age is not an issue.

At first blush, this is an attractive account of word meaning; however, even the small set of words given earlier quickly shows up its limitations. The word *bachelor* is multiply ambiguous, having at least four senses: (1) ‘unmarried man’, (2) ‘young knight’, (3) ‘person with a first university degree’ and (4) ‘young male fur seal’. The three features HUMAN, ANIMAL and MALE would be the basis for distinguishing these meanings, but they would not be enough to ‘disambiguate’ even a straightforward sentence, such as ‘The bachelor died.’ Simple stylistic differences, such as between *husband* and *hubby*, aren’t captured by this procedure either. We could add more features

– but how many more features? It is hard to imagine how this kind of mechanical approach could possibly encapsulate the remarkably expressive capabilities of a language. It is a controversial area and you can find out more by following up the texts in the further reading section.

Natural semantic metalanguage

The idea that meanings can be broken down into simpler meanings by a fixed set of features is at the basis of the Natural Semantic Metalanguage (NSM) approach. This was originated by Anna Wierzbicka and has been co-developed by her and Cliff Goddard over four decades, largely in Australia. The core NSM metalanguage consists of 65 semantic primes (e.g. *I, you, someone, people, want, do, say, know, good, bad, before* and *after*) and a simple grammar. This small but flexible metalanguage has been used to capture the complexities and intricacies of meaning, including highly cultural meanings.

The NSM approach aims to minimise the inherent danger, when attempting to study meaning, of getting ‘tangled up’ in circular and highly obscure definitions. When working across a language barrier, it provides a way to avoid the ‘invisible Anglocentrism’ that happens when words from other languages, cultures and epochs, radically different from modern Anglo English, are described through the prism of English-specific words. Those using the NSM approach study diversity and change in meaning and culture, while at the same time recognising a common basis for human conceptualisation. It has been applied to emotions, values and ideals, cultural categories for people and places, words for cataloguing ways of acting, speaking and feeling, modes of social interaction, and other areas. Although based in linguistics, the approach has been applied to adjacent fields, such as anthropology, cultural history, religious studies and philosophy.

6.5 Meaning and context

We’ve already seen examples of words that are restricted in the way they are used – *trousers* don’t happily occur alongside *countermand*, which is very particular in the kinds of nouns it allows. For this reason, dictionaries often contextualise a word when they give an account of its meaning. Words can also change their meaning depending on where they appear. A *blunt* instrument is not sharp, but a *blunt* comment may be very sharp. The colour *blue* (somewhere on the spectrum between green and violet) has a different shade or tinge depending on the noun it modifies (consider *blue veins, blue flame* and *blue haze*), and has a completely different meaning in other expressions (for example, *blue murder, blue stocking, blue movie* and *men in blue*).

The wider context is also important for determining meaning. Each of the meanings of the word *glass* in the following examples comes from the different linguistic contexts.

(6.5)

(a) She drank the wine from the **glass** [= drinking vessel]

(b) It was made of **glass** [= hard brittle substance]

(c) He looked in the **glass** and didn't like what he saw [= mirror, looking glass]

(d) The image was formed by the **glass** [= lens]

(e) Even without a **glass** we could make out the ship [= telescope, spy glass]

(f) He saw the **glass** was rising [= barometer]

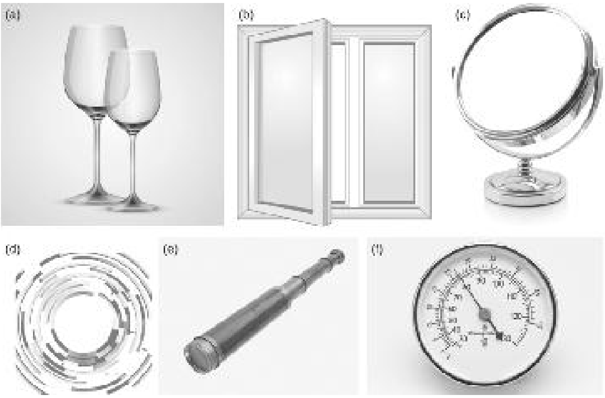


Figure 6.4 The word *glass* is polysemous and refers to all the items listed in example 6.5.

So, meaning is not just determined by the co-occurrence of words, but also stretches of text.

Clearly, all this is important when you are learning another language. We mentioned earlier that *bank* in English covers a wider range of meaning than any one word in German. It is rare to find exact equivalences, and even in closely related languages (such as English and German) there are hundreds of examples of this lack of total equivalence. German learners of English have to use context to determine whether the English word *people* corresponds to *die Menschen* (so 'people as individuals'), *die Leute* ('people as a mass') or *das Volk* ('people as a nation').

'You shall know a word by the company it keeps': Collocation

In Chapter 4, we introduced the notion of prefabricated chunks of language: sequences of words that have been stored and retrieved as single units from our memory banks. It stands to reason that the meaning (or at least part of the meaning) of individual words will derive from the words with which they habitually hang out. This is **collocation**, the common co-occurrence of certain words. As British linguist J. R. Firth once famously put it, 'you shall know a word by the company it keeps'.

For instance, the meaning of *night* is derived from constant companions like *dark* and, of course, the meaning of *dark* also derives its meaning from this close relationship. Another famous example provided by Firth is the word *ass*, meaning 'donkey' (a word that has now almost disappeared from polite conversation because of its collision with *ass* the body part, the shortening of *arse*). In Firth's time, though, sentences such as *Don't be such an ass, You silly ass* and *What an ass he is* were common. The word *ass* collocated with *Don't be such an* – and *What an* –. *Ass* also hung around with a limited set of adjectives: for example, *silly, stupid* and (one of our favourites) *egregious*.

Some words are limited in their distribution and will only collocate with certain other words. Take the adjectives *sour, rancid, rotten* and *rank* (a rare word now, but you might know it). In collocations with certain nouns, all can have the meaning 'off, bad' – but there are restrictions.

(6.6)

(a) Rotten eggs

(b) Sour milk

(c) Rancid butter (or perhaps bacon, or meat generally)

(d) Rank weeds (uncultivated vegetation emitting unpleasant odours)

So, the meanings of words will dictate how they are used, but only up to a certain point, as the examples in (6.6) illustrate, all four adjectives refer to a general state of decay, but eggs can only be rotten (hence **sour eggs, *rancid eggs* and **rank eggs* are unacceptable). Eggs can be *addled* too, if they have become rotten and failed to produce a chick. So can brains, though *addled* brains are usually ‘muddled’; the expression is used of someone who can’t think clearly. Collocations can determine particular meanings for certain words; this is another aspect of the context-dependent nature of meaning.

Frushing chickens and splatting pikes

Collocations reflect social and cultural conventions, so can change over time. These days, if you are cutting up meat at the table, you will probably use one verb, *to carve*, but in earlier times there was a flourishing lexicon for carving in English. The following are just some of these verbs (all taken from Mrs E. Smith’s *The Compleat Housewife, or, Accomplish’d Gentlewoman’s Companion*): *to frush* is what was done to a chicken, but you would *souce* a capon (a castrated cock), *unlace* a rabbit, *transon* an eel, *barb* a lobster, *splay* a bream, *disfigure* a peacock, *allay* a pheasant, *thigh* a pigeon, *splat* a pike, *mince* a plover, *chine* a salmon, *tranch* a sturgeon, *lift* a swan and *culpon* a trout. There was no difference in meaning between these verbs; it simply depended on which animal was being carved up.

Words in different languages won’t necessarily coincide in their meanings, nor do their collocations necessarily involve reliable correspondences. German has the verbs *putzen* ‘clean’ and *bürsten* ‘brush’, but beginner students of German have to learn that German people *putzen* (that is, *clean*) their teeth, not *brush* them as many English speakers do. *Putzen* can also be used in relation to noses and babies’ bottoms (not *wipe*), shoes (not *polish*) and birds preening their feathers.

Strong collocational associations between words can give seemingly neutral words a set of implied values (called their semantic prosody), which are not obvious on the surface. This is to say that the semantic values of collocates can leak back into the meanings of their partners. Consider, for example, the word pairs *largely* and *broadly*, or *little* and *small*. Each of these have unique patterns of collocation that feed back into the ways people relate to them. Corpus-based research shows that *largely* and *broadly* are clearly not the same. They cannot be substituted for each other in expressions such as *largely because, broadly speaking, largely dependent on, broadly construed, largely unaware* or *broadly agreed*.

Doing business in Russian

In a fascinating study of language change in Post-Soviet Russia, Koteyko discusses the history of the word Бизнес *biznes* 'business'. In the Soviet era, this word carried negative connotations associated with the purportedly cold and calculating ways of American business practices. *Perestroika* (the political restructuring associated with the break-up of the Soviet Union) had a limited impact on the status of this word, and Koteyko explores these developments through a pair of corpora, one constructed using texts from opposition publications and the other sourced from less overtly politically aligned publications. In opposition publications, the word *biznes* has clearly deteriorated in the post-Soviet era, becoming strongly associated with criminal activity. Koteyko shows that in a corpus of 900 000 words from the opposition press, with 395 tokens of the word *biznes*, over 75 per cent of collocates denoted illegal and unethical activities. The comparison corpus does not show this effect. The repeated use of negative collocates in the opposition publications is likely to reflect (and even foster) a negative semantic prosody for the word *biznes* among opposition press readers.

Idioms: Special examples of collocation

In Chapter 4 you were introduced to **idioms** – oft-cited examples are *spill the beans* 'to divulge a secret' and *bite the dust* 'to die'. These are complex expressions that make up single semantic units; they are extreme examples of meaning that is context-dependent. Expressions like *spill the beans* are fixed collocations whose meanings cannot be deduced from the meanings of their parts. Compare the different examples in example 6.7.

(6.7)

(a) John spilled the beans.

(b) John spilled the seeds.

(c) John let the secret out.

Certainly, one meaning of (a) is synonymous with (b) and this is the meaning you get from the composition of the meanings of its bits. But, there is another meaning (an idiomatic meaning) that is synonymous with (c). Idiomatic meanings are not compositional – you can't determine the idiomatic sense of (a) by knowing the meaning of the parts and knowing the grammatical relationships between those parts. An idiomatic meaning has to be learned as a whole (hence the label 'idiom chunks', since they are stored in one's mental lexicon as a single chunk). Note what happens when we start to move the phrases around, or alter them in any way. Straight away you get the literal sense:

(6.8)

(a) John spilled them.

(b) It was the beans John spilled.

(c) John spilled our beans.

6.6 Metonymy and metaphor

> I love metaphor. It provides two loaves where there seems to be one. Sometimes it throws in a load of fish.
>
> Bernard Malamud (1975), interview in *Paris Review* (Spring)

Definitions of **metaphor** have been many and varied over the years, and it is not our place to investigate them here. The story we tell is based on a rather general and commonsense account of the concept, but it does seem that researchers and theorists, even those of very different 'flavours', can reach agreement on this interpretation. When people use metaphor, they refer to one domain by using language expressions that normally are associated with some other domain. There is a transfer of meaning from one context to another. This explanation of metaphor has strayed little from Aristotle's original account. As he put it back in the fourth century bce, 'a good metaphor implies an intuitive perception of the similarity in dissimilars'.

Metaphor is a matter of **analogy**: speakers extend familiar concepts from their experiences and put them to new uses. Metaphor always involves the comparison of two items. Here is a straightforward example. We might want to convey a picture of someone who is totally loathsome in manner and character, so we call that person a *worm*. This sort of comparison takes salient characteristics from folk concepts about the appearance and the behaviour of the animal and these are then attributed to that person. *A worm* is 'someone who crawls and is sleazy' (in former times they were colourfully called 'lickspittles'; these days, 'brown noses' could be used). Of course, taken literally, the statement is false. This person is not actually 'a worm'; but we are claiming that there is a semantic connection between this person (the figurative meaning) and a worm (the literal meaning) – all the colour and expressive force of this insult derive from this relationship. Clearly, what metaphor does so well is to draw attention to certain features, while at the same time obscuring others. Such analogies are taken from the real world, and when these analogies are conspicuous, they are dubbed metaphors. To say *he's a worm* quite obviously applies a term from the domain of the animal kingdom to the domain of human behaviour.

The Greek philosopher Aristotle was probably the first to point to the ubiquity of metaphor. As soon as we open our mouths (or put pen to paper), we produce metaphors; however, most of them are conventionalised and we use them without conscious awareness of their relationships to more literal, original meanings. For example, in English (as in so many languages), spatial activities give rise to mental verbs, such as those of understanding, supposing and deducing. *Understand* is transparent (from the verb 'to step under'); *suppose* and *deduce* come from original Latin constructions with literal meanings, 'under + put' and 'lead + down'. Expressions to do with seeing, hearing and touching also develop to become mental verbs – for example, we have *I see* and *I hear you* for 'I understand (you)', and *He finally grasped it* for 'He finally understood'. In these cases, we are no longer conscious of the metaphorical links. Time has pushed them below the level of consciousness. Often the imagery is well and truly buried, as in the verb *comprehend*. It too comes from something that means ‘to grasp, seize’, but this is a metaphor from long ago, and one (like *suppose* and *deduce*) that has been borrowed from Latin. Much of what we talk about is in terms of something else. Scratch the surface of many expressions and you will find a dried-out metaphor of this nature.

Cooking up a dog’s breakfast and dropping off the bucket

The ‘Peanuts’ character Lucy is known for her colourful abuse, usually directed at poor old Charlie Brown: ‘You, Charlie Brown, are a foul ball in the line drive of life! You’re in the shadow of your own goal posts … you are a miscue … you are three putts on the eighteenth green … you are a seven-ten split in the tenth frame … a love set! You have dropped a rod and reel in the lake of life … you are a missed free throw, a shanked nine iron and a called third strike!’

Lucy’s tirade is a wonderful illustration of the mixed metaphor – a glorious jumble of metaphoric sporting idioms. Some mixed metaphors go unnoticed, but when the metaphorical links are the level of consciousness as in this case, the mixture of images is striking. That ‘carrot at the end of the long dark tunnel’ says it all. Such metaphoric brews might be the result of a slip of the tongue, as when Australian politician Pauline Hanson declared (in August 2018) that ‘immigration numbers are the white elephant in the room’ (or was she intentionally combining a burdensome possession that is difficult to dispose of with a very large and obvious issue that nobody wants to talk about – we’ll leave you to decide).

Linguist Karen Sullivan has investigated what happens when we mix (often incompatible) metaphors, showing how they work on a conceptual level. Her account is brimming with extraordinary images of barnacles growing on the wheels of progress, loose cannons acting as wet blankets and carrots hanging from the end of fish-hooks (when they’re not lurking at the end of long dark tunnels). But sometimes, as she shows, the mixture can be marvellous and memorable. Who could forget Texas State Treasurer Ann Richards’ description of then US presidential nominee George W. Bush as being ‘born with a silver foot in his mouth’?

One important type of metaphor is **synaesthesia**, or association of senses (from the Greek *syn* 'together' + *aisthesis* 'perception'). It can involve a transfer from sight to sound, from touch to sound, from sight and touch to intellect, and so on. The mental verbs just given are examples of this. Colours have strong psychological effects, too, and this is exploited in figurative use. Words for brightness shift to become terms to do with intellectual ability, as in describing someone as *dull* or *bright*. Colour words describing hue shift to become words expressing moods and moral qualities: *blue* has been associated with sadness since the seventeenth century and we are still *singing the blues; yellow* is associated with cowardice (but also hope), *green* with envy (or inexperience), *red* with anger, and so on.

A close relative of metaphor is **metonymy** or change by association (from the Greek *meta* 'trans' + *onoma* 'name'). A word is extended to refer to something that is closely associated with the meaning of that word. Chapter 4 gave many examples of words (such as *wellingtons* and *sandwiches*) that got their names either from their inventors or from people associated with the items in some way. *Champagne* and *cologne* are goods named after their place of origin. In restaurantese, metonymy is commonplace (staff might tell each other, *The chicken wrap wants a glass of white wine*); compare the Japanese *Watashi wa unagi da* 'I'd like an eel' (literally, 'I am [an] eel'). Metonymy is a favourite strategy for creating linguistic disguises (or **euphemisms**), so we have *spend a penny* for 'urinate' (from the days when toilets cost a penny to access), and *go to bed* and *sleep with* for 'have sex'.

An important subtype of metonymy is **synecdoche**, or the relationship of part-for-whole; something is named on the basis of its most prominent or salient part – for example, *silver* for 'cutlery' (originally made of silver) and *tit* for 'breast' (denoting the most tabooed part of the breast, the nipple). Synecdoche is often exploited in abusive language, as in calling someone a *prick*, and can be used for humorous effect – for instance, when the association is totally inappropriate, as in the Australian habit of naming people with red hair with the anomalous synecdoche *Blue*. Examples of metonymy and synecdoche differ from the earlier examples of metaphor in that they don't evoke new associations, but simply exploit those already existing between words that share a close logical relationship. Clearly, all three are very close.

6.7 Semantic change: Types of meaning shifts

Pity the poor analyst, who has to do the best he can with meanings that are as elusive as a piece of wet soap in a bathtub.

Dwight Bolinger (1975), *Aspects of Language*

The word *gay* ('homosexual') is an excellent example of a semantic shift connected to political and social change. Yet not everyone has been comfortable with this. In an essay on political correctness, Peggy Noonan (journalist and speech writer for former US president Ronald Reagan) announced that she felt the need to 'rescue' the word *gay* and return it to its 'true meaning'. This notion of a word's true meaning is a curious one. Peggy Noonan's complaint suggests she feels there is something true, natural and inevitable about her own use of language (a sentiment that lies at the heart of many people's complaints about language). The origin of *gay* can be traced back to a word meaning either 'good', 'beautiful' or 'impetuous', 'swift' (it's not clear) – but, presumably, none of these meanings is the one Peggy Noonan had in mind! 'Happy' is what she grew up with and it is this meaning she wanted to revive. Clearly, words are not those tight little bundles of form and meaning we imagine them to be, but more like Dwight Bolinger's 'pieces of wet soap in a bathtub'. In more recent times, teenagers have been taking *gay* down another quite different semantic path; driven by prejudice, *gay* has become a sneer term with roughly the meaning 'stupid'. This derogatory slang will probably not endure and who knows what meanings the word may have sprouted by the time this edition appears. One thing is certain, though, – the meaning of *gay*, like the meanings of other words, will not stand still. The recent moves to legalise gay marriage in many Western nations are no doubt having an impact on the semantic prosody of the word *gay*.

Given that there is no connection between the shape of a word and its meaning, either is free to change without affecting the other. In later chapters we will see that sounds are in a constant state of flux, but even more so are meanings. Once it was possible to *starve* from over-eating. You could eat pears from an *apple* tree, be a *meat*-eating vegetarian, drink *liquor* without touching alcohol, and be *fast* without moving at all. How is it that *black* derives from an earlier meaning 'white, shining, bright' (and is historically related to words such as *bleach, blanch* and even *bald*)? How do *grammar* and *glamour* share the same source, as do *cretin* and *Christian*?

To impose some sort of order here, semantic changes generally have been classified into various types. The classification falls along two axes. First, there are three ways that words can change in sense: meanings can **broaden**, **narrow** or **shift**. Second, there are two ways that words can change in connotation: they can **deteriorate** or **elevate**. (Clearly these descriptions are simply convenient ways to describe such changes. They have no actual explanatory value – as you might expect from their contradictory labels!) We now examine each different type of change in turn, taking examples from English and Japanese.

Broadening

Time can expand the contexts in which words appear; in short, they come to mean more than they did before. In medieval English, *dogge* referred to a particular breed of powerful canine; the word then extended its meaning to include the whole species of domesticated canines. In the 1700s, *grog* referred to 'diluted rum', but it extended to include any spirit mixed with water. From the 1800s onwards, it included 'strong drink in general'; in Australian English, it has broadened again and now refers to 'any kind of alcohol'. Other cases of broadening occur when proper names start referring to the general class of phenomena and we have many examples of this: *champagne, burgundy, cheddar, cologne, denim* and so on are all goods named after their place of origin.

The early Japanese word *ake-shida* 'dawning time' shifted to mean 'the following morning' and then broadened further to give the Modern Japanese word *ashita* 'tomorrow' (in other words, its meaning now encompasses the whole day). Many former Japanese trade names have become ordinary words; *famikon* is the name for Nintendo Entertainment System (NES), but some Japanese speakers now use the term for any video game console.

Narrowing

Time can also reduce the contexts in which words can appear: they come to mean only a part of what they originally meant. This process is more common than broadening and examples are easy to find. A *hound* was 'any sort of dog', but it narrowed to 'a particular breed of hunting dog'; *girl* described a young person of either sex, with a meaning similar to *youth* (and note, *youth* is now showing signs of restricting its application, but this time to 'young male' – when youths make an appearance in the newspaper, they are generally males). The word *liquor* once referred generally to 'beverage of any kind', but has since narrowed to 'alcoholic beverage'. Modern English's *drink* is now heading in exactly the same direction (evident in the question 'How many drinks have you had'?).

Narrowing often results when words are used in fixed situations where people share common interests and habits. Many of the terms that earmark computer-speak derive from common usage but in the context of computing their scope has reduced considerably. The verb *input* dates back to Middle English, when it had the very general meaning ‘to put on’ (in the Wyclif Bible of 1382 we can read how Paul had gathered a bundle of sticks, and ‘input’ them on the fire). *Interface* has its origins in a nineteenth-century noun generally meaning ‘a surface where two portions of matter or space meet’.

The Modern Japanese word *yakuza* ‘gangster’ (someone involved in organised crime), originally had a much broader meaning that encompassed a variety of people who might generally be classed as ‘scoundrels’ in some way – that is, gamblers, playboys, vagrants, rogues and so on. Borrowed words often undergo changes of this kind. In Japanese, *sake* has a very broad meaning that includes ‘alcoholic beverage’ generally, but in English it has narrowed to mean a particular beverage from Japan called *nihonshu*.

Shifting

Time can also totally alter the contexts in which words appear. In early English, *fond* meant ‘foolish’ (so when Juliet told Romeo ‘I am too fond’, she did not meant she was infatuated with him). With time, words can even evolve to become contranyms (words with contradictory meanings). Both the current authors use the expression *lucked out* with the older meaning of ‘lucky’; most of our students understand it to have the opposite meaning – ‘unlucky’. Here the driver of this change is perhaps the expression *luck running out* (or, as the *Urban Dictionary* suggests, *shit outta luck*).

The term *grog* mentioned above comes from the nickname for British naval officer Edward Vernon, who was known as Old Grog (the inspiration for the name was his *grogram* cloak or breeches, *grogram* being a type of cloth). The name then shifted to the mixture of rum and water, which Admiral Vernon ordered to be served out instead of neat spirit to stop sailors intoxicating themselves. The sailors, as you can imagine, were not happy and dubbed the insipid beverage *grog*.

In modern Japanese, the adverbs *hayaku, sassoku* and *sumiyaka ni* mean both ‘rapidly’ and ‘immediately’; records suggest that ‘rapidly’ was the earlier meaning, and this is supported cross-linguistically – the shift from ‘rapidly’ to ‘immediately’ is extremely common in languages around the world.

More two-faced words

As mentioned previously, meaning change can give rise to contranyms. Over time, the different senses of a word can simply drift apart; if this becomes problematic, one meaning will usually drop by the wayside. The original meaning of *fast* 'not moving at all, fixed firmly' (as in *it stuck fast)* sprouted the newer meaning 'moving quickly' (as in *she runs fast).* This newer sense is now dominant. Occasionally, opposite meanings can coexist quite peacefully, but only if they occur in such different contexts that there's no chance of a misunderstanding. To *trim* something is to cut bits off, unless it is a Christmas tree, in which case you put bits on. Do *seeded* oranges have seeds, or have they had their seeds removed? As we described at the start of section 6.5, a *blunt* instrument is dull, not sharp, but a *blunt* comment may be *quite* pointed, and can be sharp. (And what did we mean by that *quite* there – *quite* 'to the utmost degree, completely', or *quite* as in 'a little, not completely'?) The two words *spendthrift* 'extravagant person' and *thrifty* 'frugal' are clearly different expressions, but they both contain opposing meanings of *thrift* – in other words, different aspects of the condition of thriving. Something similar is happening to *sanction.* To apply *sanctions* involves measures forbidding something, but the verb *to sanction* means the opposite, 'to give permission to do something'. Both these senses have arisen out of the original meaning 'law, decree'. The words just went down different tracks. Perhaps the different grammatical contexts will ensure that both these senses continue, although the outcome is generally that speakers eventually jettison one of the uses.

Changing values: Elevation and deterioration

Through our beliefs, prejudices and experiences, words take on emotive overtones that themselves are subject to change. Words can acquire both favourable associations (= elevation), and negative or pejorative associations (= deterioration).

Elevation can occur when unpleasant overtones gradually erode away or when words take on favourable overtones. Intensifying expressions, such as *terribly* and *awfully*, have lost their original strength and now mean little more than 'very' (for example, *awfully good*). This process of enlisting horrifying words for emphasis is very widespread in the world languages – 'terrible emphasis', as it is sometimes called, may even be universal. Like the English *awesome*, the Japanese word *sugoi* can mean both 'terrible' and 'wonderful'; it is at the basis of the intensifying adverb *sugoku* ‘terribly’ (for example, *sugoku oishii* ‘really delicious’). This sort of exaggerated usage (= **hyperbole**) is a major force in semantic change worldwide. Speakers constantly endeavour to enhance expression and seek more forceful, exciting ways of saying something; and intensifying expressions such as *awful* and *awesome* are destined to become shadows of their former selves (to use another worn-out expression!).

Words can also elevate by acquiring a ‘semantic halo’ (to use C.S. Lewis’ term). In Japanese, the word *shiawase* comes from a neutral term meaning ‘a chance course of events’, but now it means ‘happiness’. Semantic halos tend to come and go, though. Take the word *politician*. As the *Oxford English Dictionary* shows, it first made its appearance with totally sinister connotations – a politician in the 1600s was ‘a crafty, cunning intriguer’ (there is an early description of the Devil as ‘so famous a politician’). It wasn’t until the nineteenth century that the word took on a more neutral sense, as the idea simply of ‘one engaged in political life; in conducting the business of the state’. It was the shift from permanent monarchies to democratically elected representative governments that helped the word to climb out of the semantic abyss. But, the word hasn’t stood still: the *Urban Dictionary* has three senses for *politician:* (1) a person who practises politics; (2) one who has perfected the art of lying; (3) a highly paid yes-man. Even mainstream dictionaries now provide tainted definitions that are approaching the earlier ‘shrewd schemer’. *Politician* seems to be coming full circle.

It is, in fact, far more usual for words to take on negative overtones rather than favourable ones. Words to do with chance are particularly prone to this. Quite different from the Japanese *shiawase*, the English noun *accident* has moved from ‘chance event’ (a meaning preserved in expressions like *by accident*) to ‘chance misfortune’, as in *there was an accident*. Australian English *mozz* (from the Hebrew *mazel* ‘luck’) shows a similar shift from ‘luck’ to ‘bad luck’, as in *put the mozz on (something)* ‘to put a jinx on something’. (As we mentioned above, it’s also not looking good for the older meaning of *luck out* ‘to be in a fortunate position’!). In Japanese, changes of this sort are also common; the word *inga* ‘bad fate’ derives from a neutral term that meant simply ‘chance course of events’. There is undoubtedly an interesting psychological basis to this sort of deterioration. Perhaps we are all inherently pessimistic. Certainly, we are far more likely to look for the worst in things. We scold and disapprove far more than we applaud and admire; bad news is always more interesting than good news. All this is reflected in the way our words change.

Why words shift their meanings

There are many reasons why words change their meanings. Earlier examples showed a close relationship between social change and semantic change. More than any other aspect of language, words and their meanings are tied to the social and cultural aspects of a speech community, and this can take them on strange journeys. Any shift in societal mores or attitudes will have semantic correlatives. Consider the meaning shifts that are currently taking place in words such as *celebrity* (‘solemn (religious) ceremony’ > ‘popular public figure’), *charisma* (‘divinely conferred power’ > ‘popular aura’), *personality* (‘pertaining to the person’ > ‘celebrity’), *image* (‘representation of person/thing’ > ‘cultivated favourable public reputation’) and *model* (‘representation of person/thing’ > ‘ideal representation’ > ‘sexualised clothes-horse’). What societal changes do you think are driving such shifts?

We have also seen examples of how prejudice can drive long-term and widespread changes, particularly affecting the value judgements attached to words associated with certain human groups. Through what amounts to little more than snobbery, words denoting a certain rank or occupation come to refer to the (real or imagined) moral or intellectual qualities of the group. Witness the elevation of terms such as *noble* and *gentle*. Not all the nobility were honourable and magnanimous, nor were all gentlewomen and gentlemen docile and courteous. Yet, these were high moral qualities conventionally attributed to people of this social status. Other examples include the elevation of *courteous* (< ‘behaviour pertaining to the court’), *generous* (< ‘of noble birth’) and *chivalrous* (< ‘characteristic of the medieval knight’). By comparison, words denoting inferior social position are quick to become morally devalued. For example, *vulgar* meant simply ‘pertaining to ordinary people’, but soon came to mean ‘offensively coarse, low’; *common* shows a similar shift.

Semantic change can also be driven by all sorts of factors arising out of our mental make-up. We’ve just seen that, in languages around the world, the meaning ‘rapidly’ typically shifts to ‘immediately’. But humans are also natural procrastinators (‘why do today what you can put off until tomorrow’) and very often, words to do with punctuality (for example, *soon, anon* [now literary], *by and by*) often end up shifting from the meaning ‘immediately’ to ‘in a little while’. *Presently* and *momentarily* are currently heading in the same direction in many English dialects. We have also seen that speakers are natural exaggerators and hyperbole is a major force behind meaning changes (an example is the use of *terribly* or *sugoku* ‘terribly’ as intensifiers). Humans can be pessimists, too, so *accident* and *inga* have shifted from ‘occurrence’ to ‘bad occurrence’. They also play down the truth in certain (often unpleasant) contexts – hence we have *pass* [*away*], or *naku naru* in Japanese ‘become non-existent’ for ‘die’; this is **understatement**.

Gresham’s Law

Taboo areas perpetually generate a narrowing and deterioration of meaning. Hard to imagine that the word *toilet* was once a sweet-smelling euphemism (from the French *toile* ‘cloth’). As the taboo sense declares itself, so the qualities of the euphemism diminish, and the next generation of speakers grows up learning it as the direct term, or worse. The result of this chronic contagion is a flourishing of vocabulary as people seek less offensive words to speak the unspeakable: *latrine* > *water closet* (*WC*) > toilet > *bathroom/washroom/restroom*. The image is that of an ever-grinding lexical mill – and, the more severe the taboo, the longer the chain of euphemistic substitutions it churns out. English has amassed more than 2500 terms for human sex organs.

Economics has Gresham’s Law: ‘Bad money drives out good.’ Sociology has Knight’s Law: ‘Bad talk drives out good.’ So why should words be any different? Linguistics has its own law of semantic change: ‘Bad connotations drive out good.’

It isn’t always the words that shift. Sometimes they stay still and the world shifts instead. We still call the traditional Christmas dessert a *plum pudding* even though we stopped making it with plums many years ago. *Pictures* once used to be painted representations, but now include paintings, drawings and photographs – in fact any image, however it is produced. The word *napkin* comes from the original meaning ‘linen cloth’, but now napkins can be made of cotton or paper. This is **subreption**. The dictionary defines it as a ‘deceptive or fallacious representation’, but in the context of language change it simply refers to the kind of meaning shift that involves the process of external change – in other words, change that occurs outside the language. The verb *write* derives from a Germanic word meaning ‘to scratch, carve’ and goes back to the time when people scratched runes into wood or rock. We continue to *write* – and sometimes we still do so with a *pen* (which originally meant ‘feather’).

Old words in a new technological world

The advent of motorisation has brought with it remarkable changes for words such as *car, tyre, lorry* and *truck*, yet the terms remain. The modern-day *dashboard* is a far cry from the board installed on coaches to stop the driver getting covered in mud and dung. *Taxis* (from *taximeter*) and *cabs* (*cabriolet*) were also very different beasts – a *taximeter* was the automatic contrivance fitted on a vehicle to indicate to the passenger at any point the distance traversed and the fare due; a *cabriolet* was a kind of public carriage with two or four wheels, drawn by one horse and seating two or four persons. Japanese has similar examples. The word *kuruma* once referred to a 'wheeled vehicle' (so included carriages and carts); in the early twentieth century, it mostly referred to a 'rickshaw' and more recently 'automobile'. *Daidokoro* has gone through various guises on its journey towards the Modern Japanese meaning 'kitchen' (in the ninth century it was a specific room in a palace where cooked food was placed onto a tray to be served to the nobility). So, it is as if the expressions outgrow their original meanings – objects, ideas and institutions alter over time, but the names for them remain.

6.8 Slips of the ear and brain

> I am sorry to say, Sir Anthony, that my affluence over my niece is very small.
>
> Mrs Malaprop (1775), *The Rivals (Richard Brinsley Sheridan)*

This section deals with slips of the ear or brain, and not 'slops of the tingue' like 'Piddle Dutch' instead of 'Middle Dutch' (an embarrassing slip made by one of the current authors during a lecture!). Slips of the ear, such as 'The winter of our disco tent' rather than Shakespeare's original 'the winter of our discontent', fall into the category of the **mondegreen** (itself a mondegreen involving the misheard lyrics of a seventeenth-century ballad: 'They have slain the Earl of Murray and Lady Mondegreen', rather than 'They have slain the Earl of Murray and laid him on the green'). Hymn-singing children come up with lots of these slips. 'Lead on Oh Kinky Turtle' is a gorgeous mishearing of the line 'Lead on, Oh King Eternal'. And then there is the cross-eyed bear called Gladly from the hymn 'Keep Thou My Way'. The original line 'Gladly the cross I'd bear' has been reshaped to 'Gladly the cross-eyed bear', something that makes far more sense to small children (as probably does 'Away in a manger, no crib for a bed; the little Malteser lay down his sweet head). There are some glorious collections of mondegreens around; just google the word and you'll find some real gems (such as the Creedence Clearwater Revival inspired 'Have You Ever Seen Lorraine').

When these sorts of mix-ups involve ordinary language, they are often **malapropisms**. This term derives from 'Mrs Malaprop', a character in Sheridan's famous eighteenth-century play *The Rivals*. Mrs Malaprop's fondness for high-falutin' sounding words led her to comic blunders such as the now famous 'the very pineapple [= pinnacle] of politeness' and 'headstrong as an allegory [alligator] on the banks of the Nile'. These sorts of slips can tell us something about the way people process language. When Australian politician Tony Abbott came out in August 2013 with, 'No one – however smart, however well-educated, however experienced – is the suppository of all wisdom', he did of course mean either 'repository' or 'depository', but while accessing his mental dictionary, something went wrong with the selection process. What has changed, of course, is the speed and ease at which such utterances reach us for derision and critique.

These days there's a new type of malapropism out there – the spell-checker malapropism (but it deserves a better label). One of our favourite examples appeared in a student's essay – in place of 'waves of righteous indignation' was 'waves of righteous indigestion'. The auto-text function on mobile phones leads to similar effects – one of us recently enjoyed an at-home *manicure* but told family and friends it had been an at-home *manure!*

It is here that we start crossing over into the territory of the newly created **eggcorn**. These are idiosyncratic slips too, but they involve homonyms or near-homonyms, and most importantly they have an inner logic to them (which Mrs Malaprops' mistakes don't really) – *coleslaw* becomes *coldslaw* (it is a type of cold salad). The term eggcorn was coined by linguists Mark Liberman and Geoff Pullum, because they felt that examples like *eggcorn* (itself an eggcorn for *acorn*) weren't handled by any of the already existing categories of language error. When a student in one of our classes came out with *rope learning* in place of *rote learning*, he probably had in mind expressions like *learn the ropes* (more meaningful for him than the rarely used *rote* 'mechanical manner').

Slips of the ear or brain can endure and bring about change. When eggcorns spread their wings and whole speech communities start using them, we move into the territory of folk or popular etymology. The line between these is admittedly rather blurred – we think of folk etymology as involving triumphant eggcorns. *Running the* *gauntlet* has nothing to do with mailed gloves like those shown in Figure 6.5. The phrase was originally *run the gatlop*, a ghastly punishment whereby prisoners ran between two rows of soldiers, forming a lane or gat, while being struck with knotted cords or sticks. *Gat* appears in place names like *Highgate*, but otherwise fell out of use. And, of course, *lope* 'run' isn't all that usual either.

Figure 6.5 A gauntlet

All these different types of reinterpretations stem from our desire to make sense of weird-looking words or expressions. It's a kind of analogy at work, involving the sorts of tidying-up activities that drive so much of language change. But, the process itself explains a lot of weirdness, too. Think of those bizarre collocations you sometimes find in pub names – *Bacchanals* probably didn't make a lot of sense in the 1700s, so people turned their local tavern into the *Bag o' Nails*.

Categories of slips aren't always clear-cut. Some years ago, one of us received a memo from the faculty office (the university shall remain unnamed), concerning a teaching committee meeting. It began 'Due to a lack of decorum'. Was this remodelling of *a lack of a quorum* a malapropism? There was perhaps a sense to it (though the meeting participants were quite well behaved!). So, it is more *eggcornish*. Then again, perhaps it was a deliberate bit of word play, so a pun all along.

And this takes us back to the mondegreen given at beginning of this section and a sign in the camping store that reads 'Now is the discount of our winter tents'!

Etymological fallacy

In June 2003, a British government minister was severely criticised for his use of the phrase *nitty-gritty* at a police conference because of its supposed racist overtones. He had told his audience that it was high time to 'get down to the nitty-gritty' about training officers. It seems that in the modern British police service at the time the expression *nitty-gritty* was prohibited because people connected it with the early slave trade. This etymology is spurious. The word entered English quite recently – sometime during the 1940s, probably via Black English and initially as a piece of popular music slang.

This is one of many bogus etymologies circulating on the internet. Normally they are a bit of fun, but occasionally they get people into trouble. Even if the etymologies were correct, word histories should not be employed in this way. The belief that a word's origin dictates what it *really* means (dubbed the 'etymological fallacy') is a piece of fuzzy thinking, especially when it is used to buttress a particular viewpoint. It was once claimed, for instance, that men couldn't be hysterical since *hysteria* comes from the Greek word for 'womb' – male hysteria was therefore a contradiction. If using etymology in this way were valid, then the chapter in this book dealing with grammar should include instruction in black magic (to reflect the fact that *grammar* once referred to 'magical lore' or 'the study of the incomprehensible' (which some might feel is still an apt description of grammar!).

The truth is, it doesn't matter what the linguistics suggests. What matters to speakers is how they perceive their language to be, and if people believe *nitty-gritty* to be racist, then this will be the kiss of death for this word. Similarly, the two nouns *fuk* 'sail' and *feck* 'purpose' have absolutely nothing to do with the infamous F-word, but that didn't save them.

6.9 Phonesthemes

Individual sounds (such as [w] and [k]) do not have meaning, and Shakespeare was correct to write 'that which we call a rose, by any other name would smell as sweet' – there is nothing to connect the form of *rose* and its meaning ‘beautiful fragrant flower’. Words are essentially arbitrary; however, the fact that people sometimes describe the word *rose* as beautiful shows that over time this arbitrariness can break down. As Chapter 1 described, speakers are always on the lookout for some sort of meaningful connection between sounds and the outside world, and they make associations between sound and sense. Recall all those *-ash* words denoting ‘violent impact’, such as *clash, crash, dash* and *flash*). The term for these meaningful sound sequences is **phonestheme**. Many of them involve the beginnings of words. The following *sn-* words convey something unpleasant (though it can be tricky to pinpoint the nastiness), and notice too that many of these words have associations with the nose: *snaffle, snout, snug, snail, snake, snare, snarl, snort, sneak, sneer, sneeze, snipe, snivel, snot, snoop, sneeze, sniff*, and so on. Phonesthemes even provide the trigger for brand-new words. *Snore*, for instance, arose as a sound symbolic word, probably inspired by *sn-* words such as Dutch *snork* ‘snore’ and *snort* (which in the 1300s also meant ‘snore’). As we saw in Chapter 4, words rarely arise out of the blue.

Phonesthemes aren’t like morphemes, but are often said to manifest a sub-regularity in the language. They are incomplete syllables, and their meanings are much more difficult to pin down. Furthermore, exactly the same cluster of sounds that constitute the phonestheme will appear in many vocabulary items that have no trace of the meaning associated with the phonestheme. You can probably think of many *sn-* words, such as *snow*, that don’t fit our earlier list at all! Just as Alice said upon hearing the poem *Jabberwocky*: ‘Somehow it seems to fill my head with ideas, only I don’t exactly know what they are.’ These are phonesthemes.

Fnasting and fneezing

Early English had the curious word *fneeze*, but by the 1500s it had changed to *sneeze* – so the fourteenth-century poet Geoffrey Chaucer’s ‘he speketh in his nose and fneseth faste’ became ‘he speaks in his nose and sneezes strongly’. True, there weren’t many *fn-* words left in English, but the fact that there were so many *sn-* words having something to do with the nose also made *sneeze* a natural fit. So, speakers remodelled *fnese* to be more like the other ‘nosey’ *sn-* words. Individual sounds aren’t supposed to have meanings – but clearly that’s not how speakers behave! Another factor assisting this change was probably the confusion surrounding the letters *s* and *f*; in early English typography, one version of *s* (the ‘long *s*’) looked very like a modern *f*, and it is feasible that the *f* in *fnese* was misread as *s* and *fnese* turned into *sneeze*, ably assisted by other members of the *sn-* phonestheme club.

Further reading

Of the many introductions to semantics available, those we recommend are Allan (1986) and (2001), Aitchison (2003), Cruse (2004), Goddard (2008, 2011), Goddard and Wierzbicka (2002), and Lyons (1995). See also the Natural Semantic Metalanguage webpage, hosted by Griffith University, for specific information on NSM. On semantic change, see Traugott and Dasher (2002). Noonan (1998) is the source piece on the changing meaning of *gay*. To find out more about the differences between *largely* and *broadly*, see Tognini-Bonelli (2001), and on *little* and *small*, see Stubbs (1995). For a study of language and politics in Post-Soviet Russia (including the meaning shifts affecting the word Бизнес *biznes* ‘business’), see Koteyko (2014). Although Bolinger (1975) and (1980) are general accounts of language that have been circulating for some time, they are wonderful books that offer an interesting take on words and meaning. Euphemism is a topic we look at later, but sources you can follow up on now are Allan and Burridge (1991) and Burridge (2012), and for metaphorical expressions we recommend Lakoff and Johnson (1981), as well as the timeless account offered by the fourth-century Greek philosopher Aristotle (see Aristotle 1984); Sullivan (2018) is an informative and hilarious account of mixed metaphors. For sound symbolism, see LaPolla (1994) and Hinton, Nichols and Ohala (1994). Benczes (2019) explores the ins and outs of phonesthemes, including how they come about, their status within morphology and their use in advertising and branding. In our discussion of the applications of semantics in a legal context, we have drawn on Cooke (1995), Eades (2012) and Koch (1985). You will find details of the now famous quotation ‘you shall know a word by the company it keeps’ in Firth (1957). The internet abounds in fascinating discussions of misheard expressions and folk etymology (especially eggcorns, mondegreens and malapropisms); the original discussion that prompted the term ‘eggcorn’ can be found on the linguistic blog ‘Language Log’ (http://itre.cis.upenn.edu/~myl/languagelog/archives/000018.html); for more political gaffes, see Manns and Burridge (2016).

Exercises and discussion points

1 Semantic components – idiomatic meaning

The following is an extract from a torrid tale of 'requited' love. Using examples, explain what aspect of English semantics and morphology this story has exploited for humour.

> It had been a rough day, so when I walked into the party I was very chalant, despite my efforts to appear gruntled and consolate. I was furling my wieldy umbrella for the coat check when I saw her standing alone in a corner. She was a descript person, a woman in a state of total array. Her hair was kempt, her clothing shevelled, and she moved in a gainly way. I wanted desperately to meet her, but I knew I'd have to make bones about it, since I was travelling cognito. Beknownst to me, the hostess, whom I could see both hide and hair of, was very proper, so it would be skin off my nose if anything bad happened … (Jack Winter, 'How I Met My Wife', *The New Yorker*, 25 July 1994)

2 Semantic connectedness – the 'nyms'

Say whether the following connected words involve a relationship of homonymy (different words with the same shape); polysemy (same word with different meanings); synonymy (sameness of meaning); antonomy (oppositeness of meaning); or hyponymy (inclusion of meaning). If the relationship is one of antonomy, then identify the type, and if it involves synonymy then try to pinpoint whether there is any difference (for example, dialectal, stylistic).

(a) thrifty, stingy

(b) site, sight

(c) weed (plant), weed (garment; for example, widow's weeds)

(d) scarlet, red

(e) stalk (part of a plant) and stalk (harass a person)

(f) old, young

(g) rancid, rotten, addled

(h) teacher, pupil

(i) fall, autumn

(j) feline, cat

3 Ambiguity – ‘crash blossoms’ in headlinese

In their effort to achieve conciseness, sub-editors often leave out the grammatical words from newspaper headlines, and the result can be amusing ambiguity. Below are some celebrated examples of ambiguous headlines, now dubbed ‘crash blossoms’ (and for more crash blossoms, see various posts on ‘Language Log’ (https://languagelog.ldc.upenn.edu/nll), where you will also discover the reason for this linguistic label).

Decide whether the ambiguity in these headlines is structural, lexical (that is, based on an ambiguous word) or both, then use paraphrases to establish the different interpretations.

(a) Truck Leads Police to Molest Suspect

(b) MacArthur Flies Back to Front

(c) City Ducks Vote on Transportation Fee

(d) Squad Helps Dog Bite Victim

(e) Stud Tires Out

(f) Red Tape Holds Up New Bridge

(g) Stolen Painting Found By Tree

(h) Kids Make Nutritious Snacks

(i) McDonald’s Fries the Holy Grail for Potato Farmers

(j) Gator Attacks Puzzle Experts

4 Semantic deterioration

Trace the history of *left, gauche* and *sinister*. Now find as many expressions as you can that include the word *left* (*a left-handed compliment*) or that refer to ‘left-handed people’. What view of left-handers is being presented here?

> By comparison, trace the history of *right, dexterity* and *ambidextrous*. What sort of associations does *right* have in expressions like *Right Reverend*? Think of others and compare these with your findings for *left*.

5 Vague resemblances – phonesthemes

Recall that phonesthemes symbolise certain meanings (although they are often difficult to pin down). Examine each group of words below, and answer the questions that follow.

(a) crash, dash, flash, smash

(b) gush, rush, smash, swish, whoosh

(c) bump, dump, hump, lump

(d) swagger, swash, sway, sweep

(e) glare, gleam, glimmer, glint

(f) grimy, grotty, grubby, greasy

(g) amble, ramble, scramble, shamble

(h) snarl, snort, sneak, sneer

(i) flitter, flow, flicker, flurry

(j) twist, twirl, tweak, twine

1 Identify the phonestheme and describe the meaning it characterises.

2 Add any other words that support this relationship between sound and sense.

3 In one paragraph, describe how you think phonesthemes might come about.

4 Give other examples of phonesthemes (in English or any other language that you know).

5 Describe how phonesthemes differ from derivational morphemes – in other words, how do they differ from conventional suffixes and prefixes?

6 Are you still happy to describe morphemes as the smallest meaningful units in language?

Research project

Semantic association

Take the list of words that follow and find ten willing participants (try to find people of roughly the same age). As you read each word out to them individually, ask them to say the first associated word that comes into their minds. Record their responses and answer the questions that follow.

(a) hamburger

(b) awesome

(c) blind

(d) carrot

(e) fire

(f) holiday

(g) picture

(h) pig

(i) poodle

(j) puppy

(k) average

(l) rich

(m) riot

(n) rotten

(o) shout

(p) smelly

(q) stinking

(r) ten

(s) terrible

(t) chair

1 Analyse each response in terms of the semantic relations (for example, synonymy, antonomy, hyponymy or collocation).

2 Take note of where people provided some sort of value judgement, and whether or not it was positive or negative (for instance, for hamburger they might have answered with *yummy/convenience* or *tasteless/overweight*).

3 Identify where there was consensus among individuals about the associations the words triggered.

4 What do you think the responses tell us about people?

5 How might advertisers or politicians – those skilled in the art of persuasion – use associative meaning and the connotative value of words?

For more activities, please go to www.cambridge.edu.au/academic/linguistics.

Part 3

◆

‘The deep grooves of language’

Sounds and grammar

7

Phonetics

◆

With acknowledgement to Jenny Price

> Take care of the sense, and the sounds will take care of themselves.
>
> Lewis Carroll, 1865, *Alice's Adventures in Wonderland*

7.0 Introduction

Phonetics is the science describing the sounds of speech. In our day-to-day lives, we don't always think a lot about how we speak. Unless there is a particular reason to focus on it, we tend to take it for granted. We might think of speech as occurring mostly between people who are in the same physical space. Indeed, until the voice could be transmitted by radio or telephone, face-to-face interaction was the only option. Today, the boundaries of space and time have exploded to the point where we can share a 'virtual' physical space with anyone in any time zone through Skype and other video chat services, including our mobile phones.

In this chapter, we will explore how speech is produced, examine its building blocks and gain some understanding of the different sub-categories of phonetics. Learning about the way speech sounds are constructed is a bit like understanding musical notes in order to play an instrument. Once you understand how the system works, it's amazing what you can do and hear!

7.1 Articulatory phonetics and the vocal tract

This section deals with the physiological aspects of speech production. Figure 7.1 shows the parts of the body used to create speech sounds – together these are known as the **vocal tract**. The vocal tract is made up of three difference spaces where air can resonate, producing sound. These spaces are in the mouth, nose and throat, and are called the **oral, nasal and pharyngeal cavities**. Sound is produced using a limited number of fixed parts of these areas. They include the alveolar ridge, upper teeth, hard and soft palate, uvula and pharyngeal wall. Also involved are more flexible parts, primarily the lips and tongue, but also the lower jaw. Together, these parts are called **articulators**. The upper articulators are fixed and passive while the lower articulators are active.

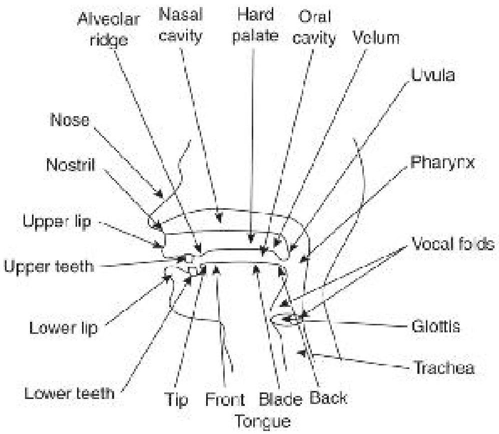


Figure 7.1 The organs of articulation

We become very conscious of our vocal tract in situations where our ability to speak is compromised for some reason. Common situations include when we are breathless from exertion, when we have had too much to drink or when we lose our voice or have a blocked nose due to illness. (A sentence like *Dean's seen nineteen underground garden gnomes in Kensington North* is a challenge for someone with a bad cold – try pinching your nose and saying it slowly to get the general idea.) Even a broken or lost tooth will make a difference to our speech, because it alters the degree of constriction we can make when producing words containing 's' sounds, for example.

To produce speech, air must be pushed from the lungs through the vocal folds inside the larynx, forcing them apart and causing them to vibrate. The frequency of this vibration, controlled by muscular tension applied to the vocal folds, is variable, and is perceived by the listener as voice **pitch**. The tighter the squeeze, the higher the sound.

Yet this is only part of the journey – once the air has passed through the 'gates' of the vocal folds, it has to navigate its way through an obstacle course of constrictions and blockages placed there by the articulators in order for the required sequence of vowels and consonants to be produced. Consonants involve a greater constriction of the airway than vowels, and the two types of sounds are described quite differently from each other. We will discuss how they are articulated in detail below, focusing on English consonants and vowels, but with occasional forays into less familiar sounds. You will notice that each sound has a characteristic symbol, which appears either in slant brackets (e.g. /p/), or occasionally in square brackets (e.g. [ʔ]); we will explain these conventions later.

7.2 Consonants

In describing consonants, we need to ask three main questions: Where in the vocal tract is the obstruction (place of articulation)? Are the vocal folds vibrating (voice and voice quality)? and, How (and to what degree) is the airflow obstructed (manner of articulation)? We begin with the different places of articulation and you will notice that the labels for the relevant vocal organs are all Latin-inspired ('labial' and 'dental' sound a lot classier than 'lippy' and 'toothy').

Places of articulation are described in terms of the articulators directly involved in producing the sound. Try making each of the following sounds yourself so you know how they feel (this works better if you keep your eyes closed). You can also benefit from looking at your face in the mirror as you pronounce them to see what your lips are doing, although you might wait until nobody else is home for this one!

- A **bilabial** consonant is one that involves both lips touching, as in /p, b, m/ in *pat, bat* and *mat.*
- The **labiodental** consonants /f/ and /v/ (*fine* and *vine*) are produced using the bottom lip and the upper teeth as articulators.
- A **dental** sound results from the tongue tip being inserted between (or just touching) the top and bottom teeth in the sounds /θ/ and /ð/ found in the words *thin* and *that.*
- **Alveolar** sounds are produced through different types of contact between the tongue and alveolar ridge (the bumpy bit behind your teeth). They include /t/, /d/, /s/, /z/, /n/, /l/ in *tack, deck, sack, zoo, now* and *loud.* If you try making these

sounds, you will be able to feel that it is only the tongue tip that is involved for /t/, /d/, /n/, whereas for /s/ and /z/ the tongue adopts something of a ‘spoon-like’ form, allowing air to pass between it and the alveolar ridge, without tongue-tip contact. The opposite occurs in the production of /l/, where the air passes over the sides of the tongue, while the tongue tip makes firm contact with the top teeth. Try saying the word *slow* a few times very slowly, to feel the different articulatory positions for /s/ and /l/.

- **Postalveolar** (or **palato-alveolar**) consonants include /ʃ/ in *fish*, /ʒ/ in *pleasure*, /tʃ/ in *church* and /dʒ/ in *judge*, all of which are produced using the tongue blade and the area of the palate located a little further back behind the alveolar ridge. Try repeating /t/ and /ʃ/ several times in succession to feel the difference in tongue position. The /ʒ/ consonant is also classed as postalveolar, and in Australian and New Zealand English, the curled tongue tip (rather than the blade or main body of the tongue) approaches this area behind the alveolar ridge. If you pronounce the /ʃ/ and /ʒ/ sounds in the words *shred* and *pleasure* slowly you can experience this difference.
- The **palatal** consonant /j/ found in words such as *young* and *yesterday* is made when the central part of the tongue simply raises and makes contact with the hard palate.
- The **velar** consonants /k/, /g/, /ŋ/ in the words *cat, got* and *tongue* are produced when the back part of the tongue lifts to make contact with the velum (soft palate).
- **Uvular** consonants, so called because the rear-most portion of the tongue is raised to touch the uvula (the dangly bit at the back that often features in cartoons of people shouting), are not characteristic of English. [R] is the symbol used to represent the trilled uvular r sound of standard German that English speakers often struggle to master, while [q] and [G] are found in Arabic.
- Finally, the sounds known as **glottals** (/h/ and [ʔ]) are less ‘consonant-like’ in that they do not use the tongue, lips or teeth. Instead, they involve the glottis (the space between the vocal folds). Although /h/ does not involve constriction or contact between articulators at any point – the vocal folds are open and the mouth is in the shape it will need for the following vowel – it is also conventionally included in the set of glottals. When the vocal folds are completely shut, the

> airstream is blocked and the glottis produces what is known as a glottal ‘stop’ [ʔ] (actually a short moment of silence, rather than a sound). Familiar examples of glottal stops exist in some English accents (for example, Cockney), in a word like *met* where the word-final /t/ is glottalised rather than pronounced (transcribed as /meʔ/). Accents with glottals have become so popular that most television series coming from the United Kingdom are full of them. How many glottal stops do you think there could be in the next sentence? *Get out at the bottom of Cotton Street, turn right and go straight for about 848 metres*.

Voicing

We now have some of the basic building blocks of sounds to work with, but a couple more things are involved. For instance, both /f/ and /v/ are classified as labiodental sounds, but we need to be able to distinguish between them. Try saying the words *fat* and *vat* aloud with your fingers touching your larynx and you’ll feel the vibrations in the throat area for the /v/ sound that are not there for /f/. The difference between the pronunciation of these words is that /v/ is **voiced**, whereas /f/ is **voiceless**. If you say ‘f-f-f-f-f-f-v-v-v-v-v-v’ a few times, you will become aware of the vocal folds making an effort to come together to make the /v/ sound, thus producing vibration.

For /f/ the vocal folds remain open and the airstream from the lungs passes unimpeded through the glottis to the teeth and lips (that is, the *place* of articulation). Voicing is thus one means of distinguishing between two sounds produced at the same place of articulation. It is important in English, where pairs of words such as *fat* and *vat* derive separate meaning simply through the voicing quality of one sound. The voicing or voicelessness can also occur at the ends of words, such as in the pairs *back* and *bag*, and *rice* and *rise*.

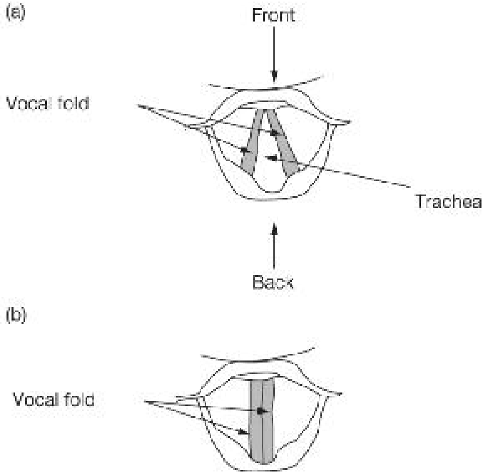


Figure 7.2 (a) Open vocal cords (b) Closed vocal cords

Voicing is far less important in other languages; for example, in many Aboriginal languages, including Gurindji, voicing is not significant and a bilabial /p/ may be more or less voiced without changing the meaning of the word.

Manner of articulation

Along with voicing, speech involves changes in how the air is pushed out of the body in a number of other ways. **Manner of articulation** depends on how the air makes its way through the vocal tract and the extent to which it is restricted. Let's start with the distinction between oral and nasal sounds. Both /b/ and /m/ are bilabial, voiced sounds, but they differ in manner. If you say /b/ but try to sustain it as long as possible before making the sound, you will feel the air pressure build up in the mouth as it seeks to burst out between the lips. If you then make a sustained /m/ sound, cease it while maintaining closed lips and repeat this a few times, you can feel that the soft palate or velum lowers to allow the air to escape through the nose to produce what is known as a **nasal** sound (or consonant). The velum raises again to close off the exit when you stop making the [m] sound. It is easy to experience the way nasal sounds involve air leaving through the nose – just say an [m] for a while and pinch your nose closed halfway through (or think of the way someone sounds when they have a blocked nose from a cold). This contrast is one of **orality** and **nasality** (with /b/ described as an **oral** sound and /m/ as a nasal sound) and, like the voicing distinction, it allows words such as *bat* and *mat* to have separate meanings in English. The only other nasal sounds in English are /n/ and /ŋ/ (the sound at the end of *sing*); as we will see in Chapter 8, /ŋ/ is restricted as to where it can appear in words.

There are many different types of oral sounds in the world's languages, and we introduce the more common ones below, along with some that are less common but fun to try. To demonstrate how air flows through the mouth, we will make use of manner diagrams. These show the way different parts of the mouth come together and separate across time. The top line represents one of the fixed articulators while the bottom line indicates the movement of the lower, active articulator. For example, in Figure 7.3, the diagram represents the closure associated with the sound [d]. The top line represents the alveolar ridge and the bottom line represents the tongue; the same figure would be appropriate for [t] as well.

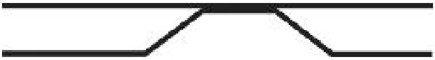


Figure 7.3 Manner diagram for [d]

Stops

Sounds that involve a complete block of the airflow are known as **stops** or, more technically, **plosives**. The English stops are /p/, /b/, /t/, /d/, /k/, /g/ and [ʔ]. The nasal consonants /m, n and ŋ/ can also be referred to as stops because they prevent air from leaving the oral cavity, even though, as described above, they allow it to escape through the nasal passages. Stops are then further categorised according to their place of articulation. Thus /p, b, m/ are bilabial, /t/, /d/, /n/ are alveolar, /k/, /g/, /ŋ/ are velar and [ʔ] is a glottal stop. The articulators responsible for stopping the airstream are the lips, tongue/alveolar ridge, tongue/velum and vocal folds/glottis respectively. There is one important point to note about the nasal sounds /m/, /n/, /ŋ/. These sounds are made using both the mouth and the nasal cavity. So, a manner diagram for an English nasal stop requires two parts – one for the oral cavity, showing the articulators coming into contact and releasing, and one for the nasal cavity, which is open. In each case, the manner diagram in Figure 7.4 applies.

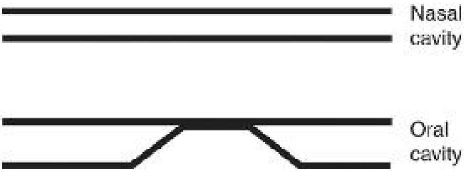


Figure 7.4 Manner diagram for nasal stops in English

Stops are labelled in contrast to a large set of other types of sounds, which are made by partly constricting the airway; together, these sounds are called **continuants**. There are two categories of continuants, **fricatives** and **approximants**.

Fricatives

Fricatives involve quite a narrow constriction of the airway that causes friction between the articulators. The word *fricative* comes from the Latin *fricāre* ‘to rub’. This results in various types of hissing sounds, their nature depending on the place of articulation. They include the following well-known sounds;

- labiodental /f/ and /v/ (lips and teeth) as in *fan* and *van*
- dental /θ/ and /ð/ (tongue and teeth) as in *thigh* and *thy*
- alveolar /s/ and /z/ (tongue and alveolar ridge) as in *sue* and *zoo*
- postalveolar /ʃ/ and /ʒ/ (tongue and postalveolar area) as in *bash* and *beige*.

The first fricative in each pair in this list is voiceless. The consonant /h/ is also included here as a voiceless fricative. In some analyses, it is regarded as a voiceless approximant (see below) but in practice voiceless approximants are rarely distinguished from voiceless fricatives.

The movement of the articulators is illustrated in Figure 7.5; as you can see in this diagram, the two articulators never completely meet.

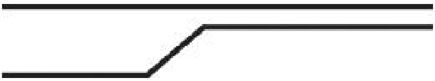


Figure 7.5 Manner diagram for a fricative

Affricates

Affricates are interesting because they combine characteristic features of both stops and fricatives. They begin with a complete closure of the airway with a gradual, fricative release (so they are rather like a stop, but with a slow release). There are only two of these sounds in English – postalveolar voiceless [tʃ] as in *church* and its voiced counterpart [dʒ] as in *judge* – but other languages have many more. See also Figure 7.6, which shows the initial complete closure followed by a partial release.

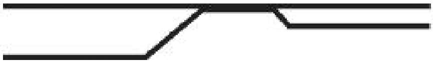


Figure 7.6 Manner diagram for an affricate

Approximants

Approximants are produced by the speech organs coming close to each other, but not close enough that an audible turbulence is produced; so, they are named after the fact that closure in the oral cavity is only **approximate**. In other words, there is still plenty of room for the air to move through the articulators. English has four of these sounds: /j/, /w/, /ɹ/ and /l/, and their articulation is indicated in Figure 7.7. They are all voiced and are found at the beginning of the words *you, woo, roo* and *loo.*

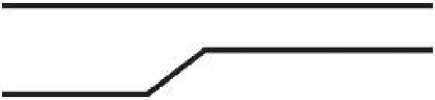


Figure 7.7 Manner diagram for an approximant

Even though the description here makes approximates appear rather laid-back little sounds, the tongue may need to be rather more gymnastic in producing them than for most other sounds. For example, the /j/ at the start of *your* (which is produced by raising the tongue to the palate directly above and 'tensing' it for a matter of microseconds) requires a quick change for the following /o/, where the tongue is back and low.

The approximant /w/ as in *way* is described as **labiovelar** because it requires movement of both the back of the tongue (in the direction of the velum) and rounding of the lips at the same time. Try saying *wheel* several times as fast as you can and you'll feel your tongue working hard to make the three sounds that make up the word – and no cheating by making the /l/ into another /w/ either (in many parts of Australia and New Zealand, word-final <l> is pronounced as [w]).

The most variable of the approximant consonants are the R-like sounds, or rhotics. This is because they can be produced in several different ways. Most common in Australian and New Zealand English is the postalveolar /ɹ/, where the tongue tip approaches the area behind the alveolar ridge in words such as *riddle* and *track*. As described earlier, [R] is the symbol used for the uvular trill made by standard German speakers, and [r] indicates the rolled or trilled *r* sound, where the tongue tip taps repeatedly against the alveolar ridge. This occurs in a variety of languages, including Mali, Russian, Italian and Spanish, some accents of French and Scottish English.

Note that the <*r*> is not pronounced when it follows a vowel in words such as *bar* and *fort*, because Australian and New Zealand English are *non-rhotic* (or R-less). Examples of *rhotic* (or R-ful) dialects are American, Canadian and Irish English, where the <r> is pronounced after a vowel (or post- vocalically). However, one context in which post-vocalic /ɹ/ *is* pronounced in non-rhotic dialects is where it precedes another vowel, such as in the phrases *for example* or *pour it out*. Here it is called a ‘linking R’. Another is when we produce an ‘r’ sound where none exists, to facilitate what would otherwise be a difficult ‘joining’ of two vowel sounds, as in *Julia and Mary*, which for most speakers of Australian and New Zealand English is pronounced *Julia-r-and Mary*. In this case it receives a more prescriptive label – the ‘intrusive R’.

Finally, the *lateral* approximant /l/ involves air escaping over the sides of the tongue while the tongue tip is making contact with the alveolar ridge. Australian and New Zealand English are two of several varieties that have both *clear* [l] and *dark* (or ‘velarised’) [ɫ] in their sound inventory, depending on where the lateral occurs within a word. We will be exploring this more in Chapter 8, but basically clear [l] is produced as just described when /l/ begins a word like *late*, whereas dark [ɫ] is characteristic of word-final /l/ in *tale*, but also in words where /l/ becomes, a stand-alone syllable in words like *table*. Dark [ɫ] is said to be velarised because, although the tongue tip still contacts the alveolar ridge, the tongue body position is more retracted and raised towards the velum. Try saying *ladle* a few times and you should feel this difference.

In sum, we can describe the consonants of English in terms of the following three parameters:

1 place of articulation

2 state of the vocal folds (voiced or voiceless)

3 manner of articulation.

It is not necessary to state specifically whether consonants are oral or nasal, given that the velum is raised for all the manners of articulation, except nasals (and nasality is already encoded in the name of this term). Likewise, voicing is only mentioned for those places and manners of articulation where a voicing contrast occurs, such as with alveolar stops. Consider the descriptions for the consonants at the start of the words in the following Mark Twain quotation: ‘**D**rawing on **m**y **f**ine **c**ommand of **l**anguage, I **s**aid **n**othing.

(7.1)

[d]	voiced alveolar stop
[m]	bilabial nasal
[f]	voiceless labiodental fricative
[k]	voiceless velar stop
[l]	lateral approximate
[s]	voiceless alveolar fricative
[n]	alveolar nasal

Ugh! Blah! Eek! Uh-oh! Phwoaaarr!

From a phonetic point of view, the class of words known as **interjections** is very interesting. Interjections are words that stand for those involuntary noisy responses we make to express some sort of emotion. They include a range of well-bred, dignified utterances such as *oh, tut-tut, tsk-tsk, ho-ho, hoorah* and *bravo* (those likely to make it onto lists of interjections in grammar books), and also the less dignified curiosities such as *ugh, blah, eek, uh-oh, meh* and *phwoaaarr* that flourish in comic books.

Clearly, outbursts such as *ugh* and *phwoaaarr* don't fit the normal pronunciation patterns of English, so these words always start life as rather inadequate attempts to represent the actual sound being made in writing. If you think about it, *ho-ho, yuk-yuk, guffaw* and *tee-hee* don't sound much like actual laughter, and *fie, faugh, bah, pish, pshaw* and *humph* aren't much like the noises we make when we are annoyed or disgusted either. But over time, as these written forms become more and more familiar, they are conventionalised as ordinary vocabulary items. So, when we encounter *ha-ha* or *haw-haw* in a text, it seems a perfectly normal representation of laughter. *Yackety-yak* seems a perfectly good way to represent someone talking. These expressions may even start functioning as ordinary nouns and verbs. Words such as *burp* and *hiccough* started life this way – written representations of our automatic noises (genteel or otherwise) that then became regular parts of speech.

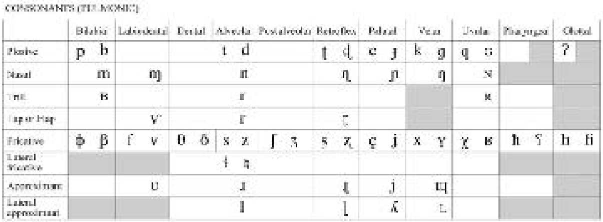
Once these expressions develop their own spelling pronunciations, their pronunciations can remove them even further from the original noises they are supposed to represent. *Tut-tut* and *tsk-tsk* don't sound much like the click we make to express impatience or disapproval (the phonetic symbol for this is [!]). Yet *tsk-tsk* now seems a reasonable way to represent this noise. Furthermore, so completely do we come to accept these as somehow natural that they then become the actual noises we make. We might say *tsk-tsk* when we are ticking someone off. Some people actually do say *yum* when eating something mouth-watering, even though this word is nothing like the sound of smacking lips. We might cry *ouch* when we bang our thumb with a hammer too (probably along with a few other less well-bred vocalisations). *Ouch* started life as a rather inadequate representation of the shriek or squeal we make when in pain.

When you consider the range of bizarre noises we release in response to emotional situations, we could never hope to represent them accurately with speech sounds. Yet, over time, these sounds and spellings somehow come to be quite appropriate.

The International Phonetic Alphabet

This problem of representing speech sounds in writing also applies to the sounds we've described for English so far. You will have noticed that while sounds such as /b/, /p/, /m/, /s/, /z/ and /w/ (among others) are written in familiar ways, other sounds are represented by different symbols – for example, /ʃ/ for the sounds spelt *sh* or sometimes *ti* (as in *nation*), and /θ/ for one of the sounds spelt *th* as in *thin*. Since each writing system is particular to its own language and since even within a single writing system there can be fairly complex relationships between letters and sounds, the International Phonetic Alphabet (IPA) is used by linguists to represent speech sounds unambiguously.

Part of the IPA is shown in Figure 7.8. It was first devised in the late 1880s and was intended to make the representative symbols for sounds more uniform across languages. It is a useful reference system when putting together the phonemic inventory (a list of all the meaningful sounds allowed) for a given language. The symbols themselves were inspired mainly by Latin letters, but also include some from Greek, such as /β/ and /γ/. Figure 7.8 shows the table of pulmonic consonants (made with air pushed up from the lungs); you'll see that they are horizontally and vertically organised in terms of the place and manner of articulation with which you are now familiar.



CONSONANTS (PULMONIC)

	Bilabial	Labiodental	Dental	Alveolar	Postalveolar	Retroflex	Palatal	Velar	Uvular	Pharyngeal	Glottal
Plosive	p b			t d		ʈ ɖ	c ɟ	k g	q ɢ		ʔ
Nasal	m	ɱ		n		ɳ	ɲ	ŋ	ɴ		
Trill	ʙ			r					ʀ		
Tap or Flap		ⱱ		ɾ		ɽ					
Fricative	ɸ β	f v	θ ð	s z	ʃ ʒ	ʂ ʐ	ç ʝ	x ɣ	χ ʁ	ħ ʕ	h ɦ
Lateral fricative				ɬ ɮ							
Approximant		ʋ		ɹ		ɻ	j	ɰ			
Lateral approximant				l		ɭ	ʎ	ʟ			

Figure 7.8 International Phonetic Alphabet (IPA) chart showing consonants
Note: Where symbols appear in pairs, the one to the right represents a voiced consonant. Shaded areas denote articulations judged impossible.

There is a lot of extra information in the full IPA chart (shown on the back cover of this book) that's not vital to your understanding of phonetics at this stage, so we'll only refer you to it where necessary. Just remember that each symbol is associated with a single speech sound, and think of the chart as a useful tool for keeping them organised. We'll return to the practicalities of using the IPA in section 7.4.

Clicks

The full IPA chart includes a series of consonants *not* produced from air being pushed through the vocal tract by the lungs (that is, non-pulmonic consonants). A sub-set of these includes the 'click' consonants that belong to the Khoisan and Bantu languages of southern Africa (represented by symbols such as [ʘ] for the bilabial click (sounds like a 'kissing' noise), [ǀ] (dental, with the blade of the tongue as the articulator), [!] (alveolar, with the tip of the tongue as the articulator), [ǂ] (palatal, using the tongue blade) and [ǁ] (lateral, using the tongue tip). Have a go at making them yourself!

7.3 Vowels

Unlike most of the consonants, vowels don't restrict the airstream through the oral cavity. In our case-study languages, vowels are voiced sounds but it is also possible to find examples of voiceless sounds in certain contexts: in Japanese, for example, some vowels are devoiced between voiceless consonants (e.g. the first vowel in the

Japanese word *sukiyaki* is voiceless, indicated here by the empty circle [su̥kijaki]). Different vowels are made by changing the shape of the mouth. Although somewhat idealised as a diagram, the cardinal vowel quadrilateral shown in Figure 7.9 represents the physiological limits of vowel realisation. Think of this chart as corresponding roughly to the area of the mouth where these sounds are produced. The labels 'front', 'central' and 'back' correspond to the parts of the tongue that are involved, and 'high', 'mid' and 'low' describe the tongue position (e.g. the vowel [i] is produced with the front of the tongue as high in the mouth as possible – a higher position would produce friction, in other words a consonant).

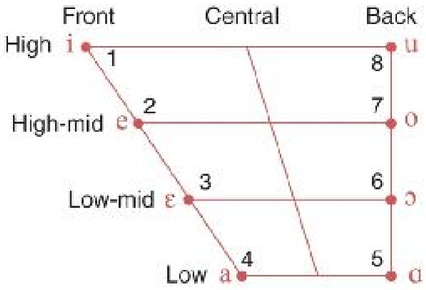


Figure 7.9 The primary cardinal vowels

'Say cheese' forces us to spread our lips into a position approximating a smile for the benefit of the photographer. 'Say aaaah' is often requested by doctors and dentists to make us open the jaw and lower the tongue body so they can see our throat or teeth. An exaggerated pronunciation of *cheese, moo* (think of the actual sound a cow makes here!) and *aaaah!* give us an idea of the outer limits of vowel production.

If you say *ee-aah-ee-aah-ee-aah* a few times (that's /i/, /ɑ/, /i/, /ɑ/, /i/, /ɑ/, in the IPA) and exaggerate the articulations you'll feel your tongue (and your jaw) moving up and down. This is because /iː/ is a high front vowel, whereas /ɑ/ is a low back vowel. The [i], [u], [a] and [ɑ] symbols represent the limits of tongue position across all languages. We cannot produce a higher or more fronted vowel than /i/, or a lower and more retracted vowel than /ɑ/, because the tongue body has reached the limit beyond which production of speech sounds becomes almost impossible.

Vowels are fascinating because they are essential to a dialect's character. For example, it is largely differences in vowel sounds that produce the distinctive accents of Englishes around the world. Think about how differently the vowel in the word *bath* is pronounced by someone from the north of England where it rhymes with the short [æ] sound in *cat*, compared with the longer, lower and more central [ɐː] of some Australian or New Zealand English speakers, and the distinctive [ɒː] variant of the South African English accent, which is produced with the tongue about as far back in the mouth as it can go, and with some lip-rounding.

Expanded versions of the IPA vowel chart (shown in Figure 7.10) can be used for describing the vowel sounds of the world's languages. It is quite remarkable that all the variability of human vowel production fits inside this four-sided shape.

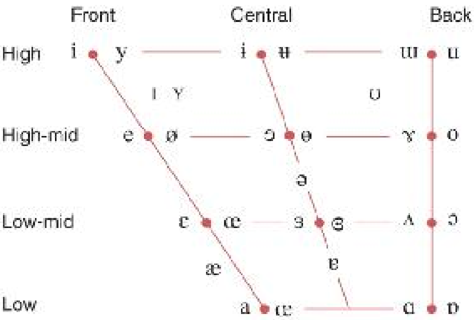


Figure 7.10 Expanded version of IPA vowel chart
Note: Where symbols appear in pairs, the one to the right represents a rounded vowel.

The basics

You can see from the full IPA vowel chart (Figure 7.10) that vowel quality is described as ranging from **high** (or **close)** to **low** (or **open)**, and from **front** to **back**. These terms all indicate the location of the highest point of the tongue body when producing the different vowels. There is a contrast between **rounded** and **unrounded** lips. You'll also notice from the chart that, for the most part, there are pairs of symbols. The symbol to the left means the lips are unrounded or spread apart, but for the one on the right they are rounded (or pursed). If you say *ee-aw-ee-aw-ee-aw-ee-aw-ee-aw* (like a siren) you can feel the tongue body shifting more from the front to the back of your mouth, and you will probably feel your lips rounding and unrounding as well. The *aw* sound in Australian English is transcribed as /oː/ and is described as a long high-mid rounded back vowel.

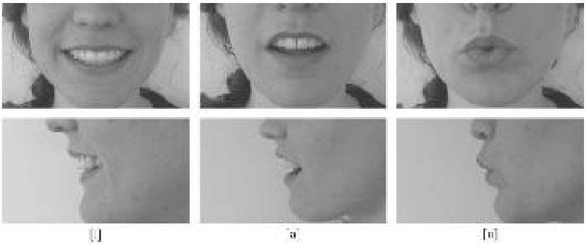


Figure 7.11 Lip positions for [i], [a] and [u]

Have a go at saying the word *cheese* and exaggerate the /iː/ vowel as much as you can. Your lips will be unrounded. Then, while keeping your tongue in the 'cheese' position, round your lips and repeat the same word. It is probably starting to sound like *choose*. This vowel /y/ has a high, fronted and rounded quality; it is not characteristic of Australian or New Zealand English, but is typical of French in the word *tu* 'you', and is written as the *ü* in German words like *Füße* 'feet'. In fact, the symbol [ʉː], for a high, central and rounded vowel, is the one now commonly used to transcribe the Australian pronunciation of the sound in *choose* (or *moo*!).

In contrast to these, right in the middle of the IPA chart is the symbol for the **schwa** /ə/ (the name schwa is the English version of the German name for a Hebrew vowel of roughly the same quality). This sound is often made with the tongue in what might be called 'neutral gear', where no articulatory effort is required – it sounds much like a relaxed version of <uh>. Schwa is the most common vowel in English and is represented by a great variety of different spellings. It appears with all the vowel letters, in fact; for example, the first vowel in words like *about* /əbæɔt/, the middle vowel in *Saturday* /sætədæɪ/ and the final vowel in *bigger* /bɪgə/. When we make this vowel sound, the centre of our tongue is slightly raised, the lips are in a neutral position (so not spread and not rounded), and there is no firm contact between rims and upper molars – this is one extremely laid-back, stress-free little vowel!

Vowels can also be short or long. The distinction is apparent in English in pairs – for example, *heed* and *hid*, or *bead* and *bid*. If you say these pairs a few times you will notice that the vowels have the same basic shape, but the long vowels take more time and effort to say. The transcriptions for each pair – [hi:d] and [hɪd], and [bi:d] and [bɪd] – make use of different vowel symbols as well as representing length. This reflects the fact that the shorter vowels are positioned a little more centrally than the related long vowels. More generally, the distinction in English is often referred to as involving **tense** and **lax** vowels, so /iː/ *heed*, /ɐː/ hard, /ɜː/ *heard*, /oː/ *hoard* and /ʉː/ *who'd* require more effort and are longer than /ɪ/ *hid*, /e/ *head*, /æ/ *had*, /ɐ/

hud, /ɔ/ *hod* and /ʊ/ *hood*.

Australian and New Zealand English

Speakers of Australian and New Zealand English are highly sensitised to differences in pronunciations between the two groups. A takeaway meal might sound like *feesh and cheeps* or *fush and chups* depending on who is saying what to whom. Figure 7.12 illustrates how the vowel systems of Australian and New Zealand English are mapped against the reference vowels of the IPA chart.

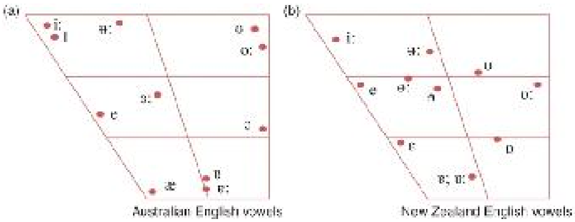


Figure 7.12 The vowel systems of Australian and New Zealand English

If you look at the reference IPA position for *hud* [ɐ] and compare it with the symbol [ɐ], which reflects how this vowel is *actually* pronounced in Australian English, you can see there is a considerable difference. (And *hud*, by the way, is either a dialectal word for the log placed at the back of a fireplace to keep the fire going, or the now obsolete word for a 'seed husk'.) It's also instructive to compare locations for sounds between the two varieties – especially [ɪ] (the vowel in both *fish* and *chips*).

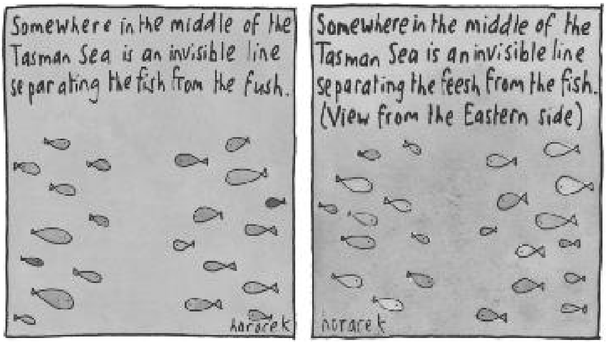

Somewhere in the middle of the Tasman Sea is an invisible line separating the fish from the fush.
horacek
Somewhere in the middle of the Tasman Sea is an invisible line separating the feesh from the fish. (View from the Eastern side)
horacek

Figure 7.13 The pronunciation of *fish* in Australian and New Zealand English

It is also a little easier to understand the characteristics of each vowel when they are displayed like this around what is often called the vowel ‘space’. High vowels are *heed, hid, who’d, hood* and *hoard*, mid vowels are *head, heard* and *hod* and low vowels are *had, hud* and *hard*. Similarly, notions of frontness/backness are more clearly identified, with *heed, hid, head* and *had* produced at the front of the vowel ‘space’, *who’d, heard, hud* and *hard* more centrally and *hood, hoard* and *hod* realised as back vowels. You can match the words to the symbols used in each case in Tables 7.1 and 7.2.

Table 7.1 Sound–spelling correspondences for Australian English vowels

Keyword	**MD (1965)**	**HCE (1997)**
heed	[illegible]	[illegible]
hid	[illegible]	[illegible]
head	[illegible]	[illegible]
had	[illegible]	[illegible]
heard	[illegible]	[illegible]
hard	[illegible]	[illegible]
hud	[illegible]	[illegible]
hod	[illegible]	[illegible]
hood	[illegible]	[illegible]
hoard	[illegible]	[illegible]
who’d	[illegible]	[illegible]

Table 7.2 Sound–spelling correspondences for New Zealand English vowels

Keyword	**Following Bauer et al. (2007)**
heed	[illegible]
hid	[illegible]
head	[illegible]
had	[illegible]
heard	[illegible]
hard	[illegible]
hud	[illegible]
hod	[illegible]
hood	[illegible]
hoard	[illegible]
who’d	[illegible]

There are two systems for referencing Australian vowels, explained further in the box on p. 194. Table 7.1 sets out the different symbols used in each transcription system. The HCE set is more fine-tuned than MD – which, to be fair, was never intended to capture more than fundamental differences. For example, the MD system does not include the length diacritic – focusing on the position of the vowel and the tense/lax contrast.

Describing Australian English vowels

Australian English was not formally described as a variety of English until the 1940s, at a time when Australia’s sense of ‘Britishness’ was still strong, and the reference accent for the ‘best’ form of spoken English for Australians was still Received Pronunciation (RP) (see also Chapter 10). Despite promoting Australian English as a legitimate variety in its own right, the scholar Alexander Mitchell described it using a set of transcription symbols corresponding to Southern Standard British English (SSBE). The MD transcription system (so called because Mitchell subsequently collaborated with Arthur Delbridge and published a set of symbols in 1965) was based on the concept of an accent ‘spectrum’ that ranged from ‘Cultivated’ (that is, most similar to SSBE) to ‘Broad’, with ‘General’ in between. While many Australian vowels shared at least some of the characteristics of their British counterparts, a few, such as /u/ in *boot*, /ɐ/ in *cup* and /a/ in *hard*, differed substantially from SSBE pronunciation, and increasingly linguists in Australia have pushed for a set of symbols reflecting greater phonetic accuracy. The MD transcription system is still in use because these symbols are universally recognised, but a parallel set was eventually developed by Harrington, Cox and Evans in 1997. Referred to as the HCE system of transcription, it is gradually becoming more widely accepted as the Australian ‘standard’. We feel you should be aware of both because you are likely to encounter both systems. The *Macquarie Dictionary*, for example, still uses the MD symbols. In this book, where it is relevant to distinguish Australian vowels, we follow the HCE system.

Other languages

Not all languages have the same number of vowels. Australian and New Zealand English have eleven, but Mali has six; Turung is similar to Mali, but has an additional three long vowels.

Table 7.3a Turung vowels

iu
ɵ, ɵɵ
a, aa

Table 7.3b Mali vowels

iu
ɘ
a

It is common to find five-vowel systems around the world and Māori is one example, with /i/, /e/, /a/, /o/, and /u/ in both short and long forms (so it has ten vowels in all). Māori has a similar system to Turung, but no central schwa vowel. Even three-vowel systems are possible – for example, as in Gurindji.

Table 7.3c A three-vowel system

i	u
a	

Diphthongs

So far, we have only discussed ‘single target’ vowels, or monophthongs. **Diphthongs** are more complex in that the tongue ‘glides’ from one position (or target) to another, demonstrating considerable articulatory range. Diphthongs are described as gliding either from a low to high position (rising diphthong), or from a high to lower position (falling or centring diphthong). Since they comprise two elements, they are represented using two symbols.

Table 7.4 Diphthongs (Australian English)

Keyword	**MD (1965)**	**HCE (1997)**
hay	æɪ	æɪ
high	ɑe	ɑe
toy	oɪ	oɪ
how	æɔ	æɔ
hoe	əʉ	əʉ
hair	eə	eː/eə
hear	ɪə	ɪə
tour	ʊə	ʉə/oː

Front-rising diphthongs in Australian English include *hay* /æɪ/, *high* /ɑe/ and *toy* /oɪ/, where the tongue position glides towards the high front region of the vowel space (see Figure 7.14). The only back-rising diphthong is *how* /æɔ/. The rising movement in *hoe* /əʉ/ is somewhere in the middle, initiating and terminating more centrally.

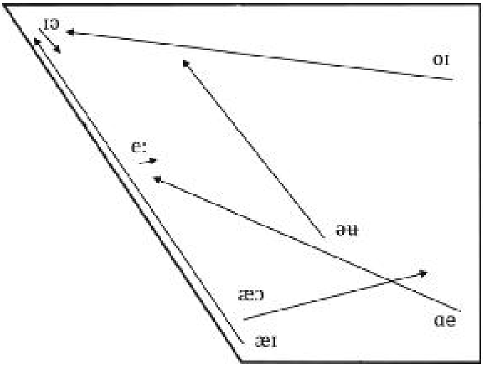


Figure 7.14 Selected diphthongs of Australian English

There are three falling or centring diphthongs in Australian English, in the words *hair* /eə/, *hear* /ʉːə/ and *tour* /ʊə/ (but their pronunciation is variable). The vowel in *hair* can also be produced as a monophthong, /eː/ (as can the one in *hear* /iː/). Just to complicate things even further, the *tour* diphthong is rarely produced by young speakers with the back /ʊ/ sound as its first element (i.e. to rhyme with words like *sewer* and *bluer*); more usually it is realised as /ʉːə/or even /oː/ (compare the pronunciation of *poor,* which also once rhymed with *sewer* and *bluer*).

In New Zealand English, front-rising diphthongs include *hay* /æe/, *high* /ɑe/ and *toy* /oe/, where the tongue position glides towards the high front region of the vowel space (see Figure 7.15). The only back-rising diphthong is *how* /æo/. The rising movement in *hoe* /ɐʉ/ is somewhere in the middle, initiating and terminating more centrally. There are also three falling or centring diphthongs, in the words *hair* /eə/, *hear* /iə/ and *tour* /ʉɐ/. Some accounts include a further diphthong /ɒɯ/ as in *gold.*

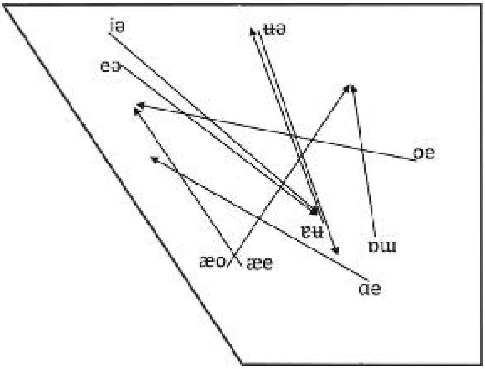


Figure 7.15 The diphthongs of New Zealand English

Table 7.5 Diphthongs (New Zealand English)

IPA (NZ)	Keyword
æe	hay
ɑe	high
oe	toy
æo	how
ɐʉ	hoe
eə	hair
iə	hear
ʉə	tour

Diphthong numbers vary across languages. Some languages have few – for example, Standard German has three (/aɪ/, /aʊ/ and /ɔʏ/) – while others, such as Gurindji, have none.

7.4 Uncomfortable relations: The written and spoken word

> Our alphabet is pure insanity. It can hardly spell any large word in the English language with any degree of certainty. Its sillinesses are quite beyond enumeration.
>
> Mark Twain, undated

As we have already hinted, the sounds of English don't match their spellings in particularly straightforward ways. Some (like Mark Twain) would go so far as to say that English is notoriously inconsistent in this regard. Painting English spelling in this way is over the top, but there are certainly some inconsistencies. The same combination of letters can often be pronounced in a variety of ways – the classic example being *through, bough, thought, rough, cough* and *though,* where the <ough> spelling is the same but the pronunciation is different for each. The corresponding IPA representations of the vowels are [ʉː], [æɔ], [oː], [ɐ], [ɔ] and [əʉ]. Conversely, the following sentence contains eleven instances of the same /iː/ sound with eleven different spellings: *f***e***verish an***ae***sthetists with am***oe***bic dysenter***y** *told p***eo***ple they bel***ie***ved they were conc***ei***ved by t***ee***nage monk***eys** *and al***i***ens wearing sn***ea***kers.*

This phenomenon occurs with consonants too. Take the consonant /t/. This sound is represented by the letter <t> as in *tap,* by double <tt> as in *butt,* <te> in *bate,* <ed> as in *jumped,* <th> as in *Thomas,* <pt> as in *receipt,* <bt> in *debt,* in *indict* – even <phth> in the word *phthisis*. In fact, there are around twelve different ways of spelling this one English consonant! The word *phthisis* is a medical term for some sort of wasting disease; it's not a term you would use terribly often, but handy if you want to point out some inconsistencies of English spelling. The <phth> is pronounced various ways: /(f)θ/, /t/, /θ/, not surprising, really, given the unusual spelling that readers are confronted with.

Sounds to symbols

Our focus here is not the vagaries of spelling, however: we just wanted to make you aware of the spelling–sound relationship because you may not have thought about it consciously. For transcription, we would like you to turn off your mental visualisation of words and simply hear their sounds. Some people find auditory tasks difficult at first, so try to concentrate on the sounds in the order they occur. A couple of tips are useful here. The letter *x* in *saxophone* is made up of two phonemes, /k/ and /s/, so it is transcribed as /ks/ – just as it sounds (whereas the sequence of letters <ph> is simply covered by /f/). And don't forget the schwa vowel. It can be represented by a variety of letters as we saw earlier and is very common.

Before we begin to transcribe sounds, it is important to know that 'phonemes' are actually abstract constructs relating to speech sounds that can have more than one pronunciation (or realisation). This is something we cover more in Chapter 8, but here is one example. The 'p' sound at the beginning of a word like *pan* is pronounced with a puff of air (known as **aspiration**). This is written as [pʰ]. It is different from the 'p' sound in something like *span*, which has no aspiration. At the end of words like *tap*, the pressure that builds up behind the lips can be either released or unreleased (and represented as [p˺]), depending on how carefully you articulate the word. So, the phoneme /p/ has at least three different pronunciations and this detail is captured in a **phonetic** transcription. A **phonemic** transcription, on the other hand, doesn't bother with this kind of information. This explains the different bracketing conventions we've been using. Phonemic transcriptions are indicated using slant brackets / … /, so all three versions of /p/ are written in the same way; however, a phonetic transcription indicates the pronunciation using square brackets and all three sounds are written differently – [pʰ], [p] and [p˺]. The bracket type makes the nature of the transcription clear to whoever might be reading it. As you may have noticed, we also occasionally use angle brackets < … > to refer specifically to the way(s) a sound is represented in writing.

In this chapter, we only go as far as phonemic (or broad) transcription so you can get used to converting the sounds to symbols. To do this, you will need to use the slant brackets. There are some excellent publications now available that cover the phonetic (or narrow) transcription of both Australian and New Zealand English, and if you are interested in this area we would encourage you to consult the books in the Further reading section at the end of this chapter.

Phonemic transcription

Let’s start with transcribing *Welcome to phonetics*. Say it to yourself and listen to the sounds. What are they? In *welcome* there’s the consonant /w/, then the vowel /e/, then another consonant, /l/ (we don’t use dark [ɫ] in phonemic transcription), then /k/ (not the letter <c>), then the schwa vowel /ə/ (it might be the letter <o> but the sound is not the one in either *comb* or *come* – if you’re having trouble with that, repeat the word *welcome* a few times and you’ll hear that you stress the first syllable but you ‘reduce’ the second). The last sound is /m/. There is no sound associated with the ‘silent e’ at the end of *welcome*, so when we put them all together we have /welkəm/. There you have your first phonemic transcription. The next word is a /t/ plus the /ʉː/ vowel, so /tʉː/, unless you reduce the vowel here, in which case it’s schwa /ə/ again. And lastly for *phonetics* we have /f/, followed by schwa /ə/, the /n/, the /e/ vowel, the /t/, the /ɪ/ or the /ɘ/ vowel, and finally the last two letters (which sound like <x>) are /k/ and /s/ – so together they form /fənetɪks/ or /fənetɘks/ depending on which side of the Tasman you’re on. The whole sentence is written between brackets as /welkəm tʉː fənetɪks/ or /welkəm tʉː fənetɘks/. See it’s /i:zi:/!

7.5 Prosody

So far we have concentrated on **segmental** phonology – in other words, information to do with vowels and consonants (the segments of spoken language). Speech also encodes a range of **suprasegmental** or **prosodic** phonology – in other words, vocal effects that spread over more than one segment. Going back to the analogy of music, prosody is largely about the rhythm, stress, pitch and tempo of utterances. In English, prosody tells us whether an utterance is a question or a statement, read text or conversational speech, the degree of familiarity speakers have with each other and even their emotional state, without us having to think about it. Stress and speech rate can also indicate the degree of formality associated with an utterance.

Syllables

Syllables are perhaps the most easily identifiable aspect of prosody. Most people intuitively know how to separate words into the small units called syllables, though they may not always agree on where to draw the syllable boundaries. This is partly because the speed of speech, degree of formality and the context of use can sometimes influence the number of syllables people assign to a word. For example, if words such as *meal* and *heal* are read aloud carefully as distinct words in a list, many people will say they contain two syllables, but when they occur in connected casual speech, they are likely to be counted as having only one syllable.

Generally, though, speakers agree on how many syllables are contained in a word – the word *syllable* has three of them! This knowledge is not the fall-out of literacy – English-speaking children can perform this task before they learn the writing system, so can speakers of unwritten languages. The scripts of some languages even have symbols based on syllables (‘syllabaries’, such as the Japanese *kana*). Moreover, as described in Chapter 8, syllables can be the basis of certain phonological processes, such as vowel harmony in some languages. Syllables are certainly real but, like the concept of the ‘word’, not easy to define. Some have tried on physiological evidence (contractions of the intercostal muscles, or chest pulses); others have used phonological criteria (based on the arrangements of vowels and consonants).

Syllables are important because they create structure in speech, and facilitate speaker expressiveness by underpinning other prosodic features such as **intonation** (the way pitch changes across an utterance), and **stress** (how loudly and how long different syllables are uttered). You are no doubt familiar with the ‘synthesised’ speech that is generated for people with speech impairments or robots in movies, and how artificial it can sound. This is partly because the rhythm, intonational contour and pitch variability of natural human speech are missing.

English uses stress to distinguish between some words. The noun *record* and the verb *record* are distinguished only by stress. In transcription, one means of marking primary stress is the diacritic /ˈ/ prior to the stressed syllable, so the noun *record* becomes /ˈɹekoːd/ and the verb *record* becomes /ɹəˈkoːd/. Similarly, compare *the tradesman was* **CON***tracted to build the house* with *the tradesman* *con***TRAC***ted an illness while building the house*. Both are verbs and the meaning of 'engaged' in the first sentence and 'developed' in the second are due entirely to the syllable bearing the stress. In English, stressed syllables tend to be slightly longer and louder than their unstressed neighbours, and their pitch is just a little higher. Test that out with the two sentences above. What happens to the vowel in the first syllable of *con***TRAC***ted*? For most speakers of Australian and New Zealand English it is reduced and realised as a schwa /ə/. This is not possible for **CON***tracted*.

The other way by which stress influences meaning is through the word in the sentence that carries the focus. *Bill found a* **PHONE** *on the beach* places the emphasis on the object, but **BILL** *found a phone on the beach* focuses on the fact that it was Bill rather than someone else who did the finding.

Languages can be described as either **stress-timed** or **syllable-timed.** Varieties such as Australian and New Zealand English are stress-timed, meaning they have a regular 'beat' or rhythm of stressed and unstressed syllables (a kind of 'tum-tee-tum-tee-tum' rhythm), whereas languages like French or Māori are syllable-timed, with virtually all syllables carrying equal stress, pitch and length (a 'rat-a-tat-a-tat' rhythm). If you are or have been a learner of French, you may have had the experience of hearing connected speech as one long stream of impenetrability. If our brain is used to the predictable rhythms of English, it can be difficult to make sense of new patterns. As an aside, a syllabic rhythm is also a feature of the English of air traffic communication ('Airspeak'), presumably because less vowel reduction comes with extra clarity.

More 'rat-a-tat-a-tat' speakers than 'tee-tum-tee-tum' speakers

The fall-out of English becoming a global language is that it has come into contact with many syllable-timed languages, and newer varieties of English (e.g. Singapore English) tend to be syllable based. So, the vowels in, say, grammatical words (like *to* or *an*), or in unstressed syllables, aren't reduced in the same way. There are even signs that syllable-based speech is becoming a marker of style in the phonological repertoires of younger New Zealand English speakers, due to the influence of Māori or the more syllable-timed Māori-influenced varieties of English. New Zealand linguists have pointed to this change in the rhythmic structure, together with the incorporation of other Polynesian features, as a significant contribution to the increasing distinctiveness of the New Zealand accent.

Structurally, syllables can be divided into four parts: the **onset**, the **nucleus**, the **coda** and the **rhyme**. Each syllable has a nucleus or peak, which is usually a vowel (V). The nucleus may have one or more consonants preceding it to create the onset and it may also have one or more consonants following it to create the coda. Syllables without codas are open, and those with codas are closed. The nucleus and the coda combine to give the rhyme of the syllable and it is this property that we make use of in rhyming slang such as *Germaine Greer* for 'beer' and *Kerry Packered* for 'knackered, tired'.

In a syllable such as *stop*, /st/ is the onset, /ɔ/ is the nucleus and /p/ is the coda, while /ɔp/ is the rhyme. We can represent its structure as shown in Figure 7.16.

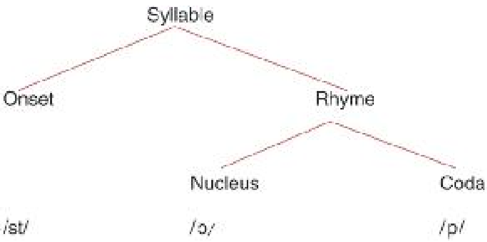


Figure 7.16 Syllable structure for stop

Note that certain consonants, known as **syllabic consonants**, can also make up a nucleus, and this is shown by a small vertical mark beneath the consonant – for example, the second /l/ in *syllable* /sɪləbl̩ / or the nasal at the end of *button* /bɐtn̩ / (some languages have a wider range of syllabic consonants).

There are languages that can have only open syllables (in other words, no codas – so all syllables end in a vowel). Māori is an open syllable language and Japanese is similar (it allows some closed syllables, but there are limits). English has an unusually large range of different types of syllables: from no codas to codas with up to four consonants, and onsets of up to three consonants, though there are restrictions (in a complex onset with three phones the first consonant must be /s/, the second must be a voiceless stop and the third a liquid, as in str*etch* or spl*at*). The monosyllabic words in Table 7.6 illustrate some possibilities for English syllables.

Table 7.6 Examples of English syllables

Syllable type	**Example**	**Symbols**
V	*eye*	/aɪ/
VC	*up*	/ʌp/
CV	*see*	/siː/
CVC	*cup*	/kʌp/

[illegible]
[illegible]
[illegible]
[illegible]
[illegible]
[illegible]

Pitch and intonation

Pitch is another aspect of prosody about which you're probably more aware than you realise. Male voice pitch is lower than that of women and children because men's vocal folds are longer and thicker, and vibrate more slowly, producing lower 'notes'. We can identify people we know over the phone from a combination of their pitch range, **intonation** patterns and voice quality, because these characteristics are already familiar to us. At the other end of the scale, speech that *lacks* variable pitch and predictable intonational 'tunes' (such as that of some deaf people) is noticeable for these reasons.

We also know that, although perhaps not entirely consciously, meaning is transmitted through the intonation 'contour' of an utterance, together with the syllables that are stressed. The question *Did you* **REA***lly*? can be interpreted as expressing surprise or disbelief, and its intonational tune (in Australian English at least) is most likely to be rising. A falling tune usually indicates a statement where there is no suggestion of doubt, such as *I'm going to* **SYD***ney next week* ↓. Sentences can be constructed of more than one intonational phrase, such as *We've been there be***FORE** ↓; *the one on the corner of George and* **CAR***ter Streets* ↓; *in the* **CI***ty* ↓. If the speaker is certain that the listener is familiar with the location they are describing, all three intonational contours would be falling. But, if the listener looks blank, the speaker might shift to a rising tune for the second and third intonational phrases, as in *We've been there be***FORE** ↓; *the one on the corner of George and* **CAR***ter Streets* ↑; *in the* **CI***ty* ↑. This rising tune is referred to as either the high rising tone/terminal (or HRT), Uptalk or Australian Questioning Intonation (AQI) (even though it made its first appearance in New Zealand English), and it can have a variety of discourse functions. Here it is used to check whether the listener knows the location they are describing. It can also be used by speakers to indicate they have not finished their speaking 'turn' – it is very common in storytelling. Although it was initially associated with younger speakers (particularly young women), it can now be heard right across the age and gender spectrum. We revisit this topic in Chapter 10.

Intonation reflects the speech style that is appropriate to the utterance and the communicative context, particularly the degree of formality and/or authority. News reading is a good example of speech that needs to be credible and authoritative, at the same time tailoring its style to the target audience of the media institution it represents. The 'older style' AM radio stations are aiming for listeners aged 50 and above, whereas FM radio is often after a much younger target market, and its speech style incorporates features that are likely to resonate with that age group, including the HRT. The dark pitch trace in Figure 7.17 shows what an utterance-final HRT looks like in the sentence *The mentally ill man who burnt down the St Kilda kiosk last year's been jailed for three years*, read by a female newsreader on a mainstream Melbourne FM radio station (you can listen to the sound file on the website accompanying this text).

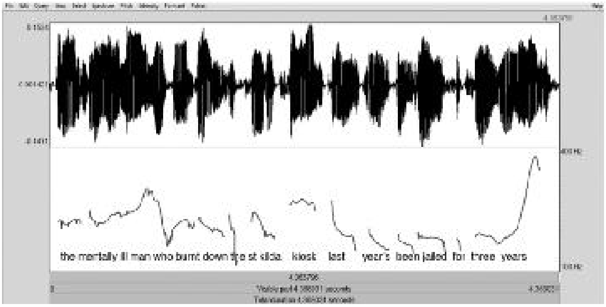

the mentally ill man who burnt down the st kilda kiosk last year's been jailed for three years

Figure 7.17 Pitch trace produced using Praat

More complex rise-fall, fall-rise and even rise-fall-rise tune types are common in conversational speech, and they are a particularly useful and efficient means of communicating attitudinal information to a listener. The exclamation *Duh!* ↑↓↑ uttered with a rise-fall-rise tune is actually saying *Of course I know that, you idiot, how stupid do you think I am!* What kind of tune describes the utterance *Whatever!*? And how would you describe its communicative purpose and the attitude it expresses?

Tone

In **tone languages,** differences of meaning contained in segmental contrasts in languages like English (for example, *pet* versus *bet*) can be communicated through the type of pitch (or tone) carried by vowels or syllables. For example, /zì/ 'self ' and /zǐ/ 'son' are two completely different words in Mandarin Chinese, with only the tone to distinguish them from each other. In Chapter 8, you will see the details of our case-study language Turung, which has three distinct tones. Given the dominance of English in the information-driven world we inhabit today, it is easy to forget how many people speak languages that make lexical and grammatical meaning distinctions based on tone (Chinese varieties and Vietnamese come to mind, but there are also the millions of speakers of African and South American languages). In terms of sheer speaker numbers this makes them the majority.

Tone languages can be defined according to their characteristic tonal systems. **Register** systems have high, mid or low tones that maintain a constant pitch across the entire vowel or syllable. Here it is the relative difference between the tones that counts, rather than the actual pitch. In **contour** systems, the tone varies through the course of the syllable, so it may fall from high to low, glide from low to high, rise from a high pitch, and so on. In these systems, it is the change in tone within the syllable that carries meaning.

Just to complicate things, some tone languages can have both register and contour tones (Thai has five tones: three register tones and two contour tones). Finally, there can be disagreement among researchers as to the categorisation itself. The examples from Vietnamese given in Table 7.7 use diacritics for tonal transcription, but some languages use numbers from 1 to 5 identifying tone height, or iconic 'tone letters' like ˥˩, which in Mandarin indicates the high falling tone (Tone 4, also represented with diacritics, as in /zì/ 'self '). Vietnamese, with six tones, belongs to the contour-tone language category. The same CV (consonant-vowel) sequence *ma* has six different meanings, depending on the 'shape' of the tone it carries.

Table 7.7 Vietnamese contour-tone

	Tone	**Example** (glossed)
(unmarked)	Level tone	ma 'ghost'
´	Rising tone	má 'mother'
`	Falling tone	mà 'which'
̉	Falling rising tone	mả 'tomb'
˜	High rising tone	mã 'horse'
̣	Low constricted tone	mạ 'rice seedling'

Yoruba (a Niger-Congo language spoken in the west of Nigeria) is a register-tone language, where lexical difference is indicated by the high, mid or low (level) tone carried by the syllable, as shown in example 7.1.

(7.2)

bí 'give birth'

bio ask’
bìomit’

If you are not a speaker of this tone language, you would want to practise a lot before being brave enough to use the ‘ask’ vowel in public!

Paralinguistic features

You may be able to think of a voice you either love or can’t stand, simply because of its inherent **quality**. Just as certain types of music can profoundly affect us, some ancillary or **paralinguistic** elements of speech (particularly voice quality, but also pitch) can provoke strong reactions. High-pitched voices might come across as irritating, or ‘breathy’ voices in women can be experienced as sexually attractive (going back in time, think Marilyn Monroe). Paralinguistic characteristics may be innate, or manifest as the short-term outcome of our emotional state. They may also be stylistically cultivated as part of our identity. Pitch is a good example: it has been recognised for decades that the pitch of young women has gradually been lowering.

A more recent arrival in female speech is ‘vocal fry’ or ‘creaky voice’, which involves very slow and often irregular vibration of the vocal folds at the bottom of the pitch register. (There are lots of wonderful examples online; just search for ‘vocal fry’.) Like many stylistic trends, vocal fry seems to be spreading fast, most noticeably in young women, and like so many other elements of youth speech (particularly of young women), it is a common target for criticism.

7.6 Unpacking the parts

Now that we have all this knowledge at our disposal, what can we do with it? So far, our discussion has focused on the individual sounds, showing how sentences would be said ‘sound by sound’ and ‘word by word’. However, all sorts of things happen to these sounds in fast speech. Take something as simple as the word *CUP*. We all think of this as being made up of three distinct sounds, /k/, /ɐ/, and /p/, represented by the letters <c>, <u> and <p>. This view of speech as a sequence of single sounds broken into words is further reinforced by writing – the fact that language is represented as a sequence of letters, separated by different amounts of space. But, this is all an illusion. Spoken language is actually nothing like this. We don’t make a /k/, then pause to make an /ɐ/ and then a /p/. In fact, if we did pronounce each of these sounds individually, and then splice them together, it would sound as if we were pronouncing the word in a very odd sort of way. And if we were to reverse the order of these sounds as we would normally pronounce them on a recording device, then the word we get won’t sound like *puck*. Sounds will always change their shape depending on where they occur in a word and what they are sitting next to.

Likewise, when we speak, we don’t just speak in single words but in groups of words. These groups are continuous, and they may or may not have pauses between them. How we differentiate words from each other is also based on an illusion. One word simply merges into the next. There are no convenient gaps between spoken words, equivalent to the little white spaces between written words, and it is astonishing how in this unbroken stream of sound we manage to distinguish words at all (recall from Chapter 4 that Steven Pinker described it as ‘hallucinating’ word boundaries whenever we reach the edge of a stretch of sound that matches up with an entry in our mental dictionaries). The (semi-)continuous flow of sounds and words also changes the shapes of the words. Some sounds might drop off, while others might be added. Some might even blend into nearby sounds, creating transitions from one sound to another. There are a number of different processes that commonly affect sounds in connected speech, and in Chapter 8 we consider in detail some of the most important of these.

Because of the seamlessness of speech, people occasionally end up identifying boundaries in the wrong places. For example, some stretches of sounds can be sliced up into words in different ways. You’ve probably encountered examples such as *ice-cream* versus *I scream; Iced ink* versus *I stink; The stuffy nose can lead to problems* versus *The stuff he knows can lead to problems; euthanasia* versus *youth in Asia*. Pairs such as these are called **oronyms** (or more appropriately ‘slice-o-nyms’). And as you might expect, people play with this particular feature of the language. You can now buy t-shirts that read ‘Be alert – the world needs more lerts’ (a quote that has been attributed to Woody Allen).

Some lovely examples appear in a skit known as ‘The Four Candles’ (written by British comedians Ronnie Barker and Ronnie Corbett, or ‘The Two Ronnies’). A scruffy-looking Barker comes into an old-fashioned hardware store and asks for ‘fork ’andles’ (handles for garden forks), and the storekeeper Corbett presents him with four candles. Corbett continues to misunderstand the items Barker reads from his list, and the sketch is full of miscommunication arising from oronyms and other kinds of mishearings. Wrong boundary assignment can endure and lead to significant changes in the language. But that is a whole ’nother story.

More mis-analysis: The mondegreen

In Chapter 6, we encountered the *mondegreen*, a useful concept when it comes to identifying the various sound segments that make up speech. Recall that some of the best examples of mondegreens occur in song lyrics. For example, in The Beatles’ song ‘Lucy in the Sky with Diamonds’, *The girl with kaleidoscope eyes* becomes *The girl with colitis goes by*.

How do we mishear these lyrics as *The girl with colitis goes by*? Our brain seeks to unravel the jumble of sound segments it hears according to what it has been trained to understand. Everything is fine up until the second syllable of *kaleidoscope*, but from there our innate knowledge of the rules of English takes over. In connected speech, the stressed /ɑe/ vowel in the second syllable of *kaleidoscope* is slightly longer than it is in the second syllable of *colitis* because the following consonant is voiced /d/ rather than voiceless /t/, but here the musical rhythm dictates phrases of equal length, which means this contrast is lost. Now, we have to get creative and the <d> of *kaleidoscope* is heard as a tapped /t/ consonant (sounding like /d/) that is common in American English, followed by our versatile friend the schwa vowel that can just as easily represent the second <i> of *colitis* as the first <o> of *kaleidoscope*. Once the word *colitis* has been established, we need to decipher what follows. The ‘unaspirated’ voiceless velar stop /k/ following /s/ in *kaleidoscope* is easily interpreted as the voiced equivalent, /g/. The <s> from the end of the word *eyes* has been transposed to a position immediately preceding the voiceless bilabial /p/ of *kaleidoscope*, which in turn is transformed into, and heard, as a voiced bilabial /b/. The final /ɑe/ vowel remains, at last allowing the brain to make sense of the whole – however ridiculous the end result might seem.

Hearing things that aren’t really there even extends to random stimuli. Some of you might recall the scandal of McDonald’s swearing Minion toys. They were speaking Minionese (nonsense prattle), but outraged parents heard swearwords: ‘Well I’ll be damned’ and ‘What the fuck’. Hundreds of others who clicked on the link heard them too. And the more they listened, the clearer these obscenities became. It’s a kind of auditory inkblot. People hear meaningful patterns in meaningless bursts of noise, just as they perceive meaningful images in meaningless blobs of ink. And once told what to listen for, or to look out for, then the brain is primed – and that’s precisely what they recognise. The fact that parents heard swear words demonstrates the saliency of these expressions (it's not for nothing they're described as 'strong language'). In Chapter 16, we'll show how all sorts of studies (linguistic, psychological, physiological and neurological) reveal that swear words are more arousing and more memorable than other language stimuli. So, if our brains are searching for patterns in random sounds, it isn't too remarkable we'll hear obscenities.

The dark side of the mondegreen

Speech slips can be hilarious, but they also have a darker side, as forensic phonetician Helen Fraser reveals. These days, many criminal trials feature evidence in the form of covert recordings (conversations lawfully captured by police without the knowledge of the speakers). These recordings can be of extremely poor quality – and, like incomprehensible song lyrics, it can often be hard to make out what is being said. What if a court then 'mondegreens' an innocent remark into an admission of guilt? Sounds far-fetched? Helen Fraser has shown that this has happened numerous times.

A judicial ruling made in 1987 allows indistinct covert recordings to be accompanied by police transcripts. The role of detectives providing these transcripts is that of the so-called 'ad hoc expert' – their 'specialised knowledge' comes from familiarity with the case and the knowledge that emerges from repeated exposure to the recording. Of course, there are all sorts of safeguards in place. The judge is expected to listen personally to the audio and make a ruling, and the jury is encouraged to listen carefully too – the transcript is only ever there to lend a hand. But there is a problem: listeners' unconscious expectations play an enormous role in shaping speech perceptions. It's simple priming – context and our own knowledge help our brains to make sense of the noise we hear. The background knowledge of the detectives investigating the case 'assists' them with their transcripts of the audio. The transcripts then 'assist' the lawyers, judges and juries to interpret the audio. Everyone believes that they are listening with their own ears, so mondegreens go unnoticed. And the consequences can be catastrophic.

At Forensic Phonetics Australia, you can hear an excerpt from a recording used in a murder trial – a whispered conversation evidently about a pact related to a murder. But the audio is indecipherable, so the judge allowed a police transcript to help the jury out. The defendant was subsequently found guilty of murder and sentenced to 30 years in prison, largely on the basis that this covert recording revealed him as saying ‘at the start we made a pact’.

It will horrify you to learn that Heler Fraser’s phonetic analysis clarifies that these words were not actually spoken, and as her later experimental studies revealed, listeners given this same excerpt ‘cold’ never hear anything about a pact at all – that is, until they are shown the transcript. One person’s mondegreen can be another person’s jail sentence.

Phonetics and sound symbolism

Phonetics can be fun and, of course, it is interesting, but it can also be lucrative. When it comes to the world of product branding, marketing and advertising, the saying ‘Knowledge is power’ might be expanded to ‘Phonetic knowledge is power’. New products have to be named before being ‘launched’ in the marketplace. This is big business, as attested by the 22.3 million results from one internet search engine when the term ‘branding strategy’ is entered. One of the secrets of successful branding relies on an understanding of the fundamental phonetic principles we have covered so far.

Phonetic symbolism, frequently referred to as ‘sound symbolism’, concerns the relationship between sound and meaning. For example, studies have consistently shown that people associate high front vowels with qualities such as smallness (think of words like *teeny weeny, itsy bitsy* and *wee*) and femininity (think of stereotypically feminine *Mimi* and *Fifi*, and diminutive endings in pet names like *Shelly* and *Katie*). On the other hand, back vowels are associated with increased size and power (‘many a mickle makes a muckle’ and while you mightn’t know what a mickle is, you probably feel it’s smaller and less impressive than a muckle!). Likewise, it doesn’t take a degree in linguistics to figure out which words in the Table 7.8 mean ‘big’ and which mean ‘small’.

Table 7.8 Words for ‘big’ and ‘small’

[illegible]	[illegible]	(Hungarian)
maliit	*malaki*	(Tagalog)
mikros	*makros*	(Greek)
petit	*grand*	(French)
kecil	*besar*	(Malay)
kidogo	*kubwa*	(Swahili)

Similarly, brand names starting with the stop consonants /p/, /t/, /k/ result in greater product recognition by consumers (Pepsi and Coca-Cola being two examples). Put simply, if the name of a product enhances its positive attributes, it will aid awareness and recall in the mind of the customer, thereby maximising the potential for sales. In an increasingly competitive marketplace, an understanding of sound symbolism can lead to success or failure. An often-cited example of a product whose spectacular failure was partly attributed to a poor choice of name is the Ford Edsel motor car, which was launched in the mid-1950s. After an exhaustive twelve months of research during which 18 000 suggestions were surveyed and tested, 'Corsair' turned out to be the most popular, yet at committee level this was rejected in favour of 'Edsel', the name of one of the Ford family members. To those who did not know who Edsel Ford was, the name simply sounded 'funny', and feedback obtained from people in the street showed they associated it with 'pretzel' and 'weasel'. While there were several factors that contributed to the car's demise, the name only added insult to injury.

7.7 Acoustic phonetics

So far we have focused on the articulatory and auditory approaches to speech analysis – in other words, the way speech sounds are produced (by the vocal organs) and perceived (by the ear and brain). We will now move on to the computer-based **acoustic analysis** of the speech signal itself. This doesn't mean that we are dispensing with our ears, however. Acoustic analysis simply allows us to 'see' (and scientifically measure) what we have already heard and assessed. Photographers often speak of the thrill they feel when they develop their images in the darkroom and, like magic, the picture appears on the paper. It's not uncommon for phoneticians to talk about a 'light bulb moment', when, as students, they first turned sound into something they could see, in a **spectrogram**. The visual representation of speech sounds has now been made very easy by freely available software programs such as Audacity and Praat.

Visualising sound

When the vibrations produced by the vocal folds are 'shaped' in the vocal tract, the resulting acoustic energy produces a recognisable pattern of resonances (that is, vibrations of air movement in the vocal tract). These appear as patterns of lights and dark, and are referred to as **formant frequencies**. While a given vowel may produce several formants (labelled F1, F2, and so on), the first three (that is, F1, F2 and F3) are usually sufficient for listeners to differentiate vowel quality. Look at the two spectrograms illustrating the words *head* (Figure 7.18) and *hud* (Figure 7.19). The dark black bands are the formants, starting with F1 at the bottom. We can see that the main difference between these two vowels is the position of F2, which in *head* is closer to F3, and in *hud* is closer to F1.

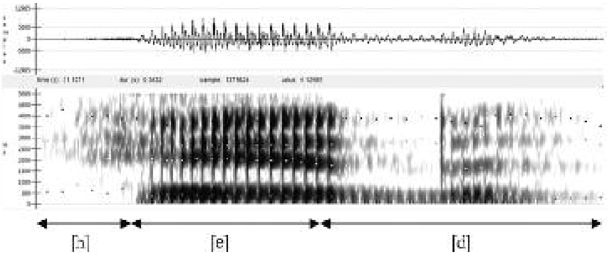


Figure 7.18 Spectrogram of *head*

Although individuals may vary considerably, these formant patterns are basically consistent across speakers, reflecting the perceptual distinction between each different vowel, necessary to avoid confusion for the listener. The spectrogram in Figure 7.20 shows the formants produced by the vocal tract of a 45-year-old, General-accented male speaker of Australian English uttering the word *Monday*.

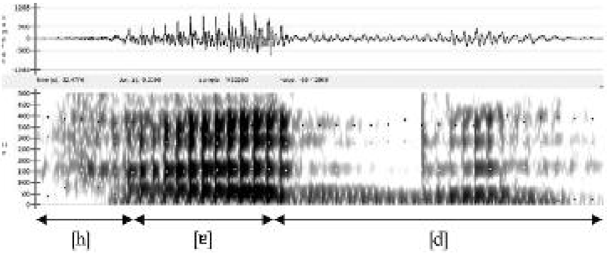


Figure 7.19 Spectrogram of *hud*

In Figure 7.20, we can also see the formant patterns for the two nasal consonants ([m] and [n]). These are less distinct, due to the reduced energy level produced by the structure and particular resonance effects of the nasal cavity. Often the energy of the first formant is easily identifiable, but subsequent resonances become increasingly weaker. For the vowels, the short [ɐ] monophthong and longer [æɪ] diphthong are clearly distinguished from their neighbouring nasal and consonant segments. Between the end of the nasal [n] and the start of the [æɪ] diphthong, we see the short burst of the [d] consonant release that precedes the onset of voicing in [æɪ]. This appears as a white area where nothing much seems to be happening, but there is in fact quite a lot going on in terms of the articulators. Nasal resonance stops and the formant structure disappears, as the velum raises to close off the nasal cavity and the tongue tip (already in contact with the alveolar ridge) remains in position ready to block the flow of air through the oral cavity prior to the release of the [d]. We can also see the glide from the first to the second ‘target’ of [æɪ] in the upward curve of F2.

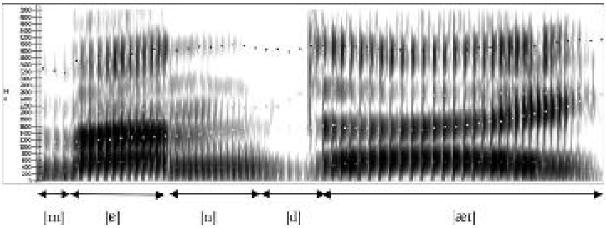


Figure 7.20 Spectrogram of *Monday*

Figure 7.21 illustrates several simple vowels (and the superimposed /æɪ/ diphthong) uttered by a group of middle-aged male speakers of Australian English. This chart is produced by labelling the mid-point of the first and second formant frequencies of every sample of each vowel on the spectrograms, then calculating the values in Hertz (the measure of frequency, shown on the left-hand side of the spectrograms above). When F1 and F2 are plotted together for each vowel, you can ‘join up the dots’ and they form a shape (often referred to as a vowel chart) that corresponds approximately with tongue-body height and fronting. The rising /æɪ/ diphthong (also calculated after labelling the spectrograms) is typical of General-accented speakers of Australian English, with the ‘elbow’ of its first element equidistant between the /æ/ and /ɐː/ vowels. The ‘gliding’ movement from the first to the second target illustrates the dynamic nature of diphthongs.

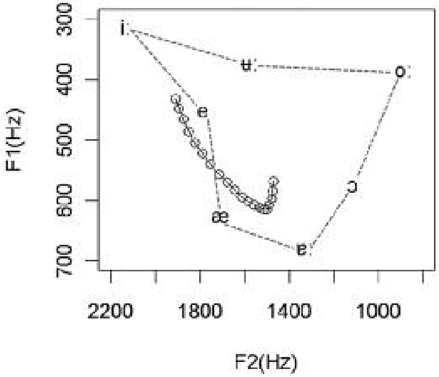


Figure 7.21 Some vowels of middle-aged male speakers of Australian English

Acoustic analysis: How it is used

One criticism commonly levelled at the articulatory and auditory approaches to phonetics is their lack of scientific rigour and objectivity. Something acoustic analysis does brilliantly is facilitate the objective comparison of data samples, whether of an individual or group of individuals, with a reference sample (such as a certain social group, or a group matched for age or gender), or samples from the same individual at different moments in time. We will look at two areas of phonetics where comparisons like these are the focus – forensic phonetics and sociophonetics – and then at the broader phenomenon of sound change. For additional applications (speech recognition and speech synthesis), you will have to wait until Chapter 18.

Forensic phonetics

Earlier, we touched on the work of forensic phonetician Helen Fraser with regard to problems arising from the use of covert recordings as evidence in court proceedings. The field of forensic phonetics is a branch of forensic linguistics, itself neatly situated at the intersection of language and the law. This is an activity that holds great appeal, no doubt fuelled by endless episodes of television police dramas. Programs like these stretch the truth when it comes to speaker identification, but let's face it – their writers have to help the good guys catch the bad guys. In the real world, forensic phonetics is a serious business, and expert phoneticians are called upon to answer questions such as the following (from actual cases).

- Is the voice of the suspect the same as the voice making the

000 call?

- Did the brothel commissioner say *I wouldn't take a bribe* or *I would take a bribe*?
- From which part of the country is the person who made the anonymous blackmail call?
- Was the captain drunk while talking to the coast guard after he had run the ship into the bridge?

While it would be very convenient, there is no 'linguistic fingerprint' or 'voiceprint' that uniquely identifies speakers in the same way that fingerprints do, so the forensic expert must find idiosyncratic features that are present in both the samples that differentiate the suspect from all other speakers. Three acoustically determined variables frequently studied in forensic work are pitch (F0) and the first two formant frequencies (F1 and F2). However, we must stress that acoustic analysis in no way replaces the auditory approach, which is still considered a necessary element of forensic work.

A few forensic cases are well known, so we'll briefly mention a couple of them. One concerns a bomb threat made to PanAm in 1984 by a man with an East Coast American accent. An employee of the company accused of the crime was acquitted on the combined auditory-acoustic evidence of linguist William Labov. He showed that the employee had a New York accent, as opposed to the caller, whose accent revealed that he was from Boston (in the quote we included in Chapter 2, Labov wrote that separating fact from fiction in this case was one of the most fulfilling moments in his linguistic career). Another case turned on a single word: whether a doctor had said *can* or *can't*. He was accused of telling a drug-addicted patient he could crush and inject the pills he was prescribing him. The doctor spoke English with a marked Greek accent, and forensic phonetician Peter French's auditory analysis showed no trace of a /t/ at the end of the word. However, his acoustic analysis of the first and second formant frequencies clearly differentiated the vowels in the *can* and *can't* words, and confirmed the doctor's claim that he'd said *You can't inject those things*.

Profile of Debbie Loakes

Debbie Loakes is a researcher at The University of Melbourne who analysed identical twins' speech for her PhD thesis using the techniques of forensic speaker identification. Since then, she has been involved in a number of forensic cases, some involving hoax calls. In one of them a male with an unusually high voice pitch was accused of making bomb threats, and Debbie's analyses included measuring the mean F0 of the forensic and suspect samples, and comparing these with the values extracted from an additional eight male speakers of Australian English. Results showed that the mean pitch range of the forensic and suspect samples overlapped, and also fell outside the population mean represented by the extra males. On its own, this analysis was not sufficient to suggest that the two samples were from the same individual, and was just one of several analyses conducted. Another case involved a caller repeatedly contacting 000 to report non-existent fires.

Some of Debbie's work has concerned recordings made with 'bugging devices', where people are either actually committing or discussing crimes. Debbie typically analyses F1/F2 to examine vowel quality, but because Australian English is a relatively homogeneous dialect, this is a less robust measure of speaker difference than it is in British or American English, for example, where regional accents are more clearly identifiable. She therefore also looks for any consistent speech habits that allow a suspect to be distinguished from other speakers who otherwise sound similar to the suspect. For instance, evidence in both samples of what is known as 'fricated /t/' – that is, a 't' that has a characteristic fricative (or 's'-like) quality – might provide a stronger indication that the speaker in the forensic and police interview recordings is the same person.

Sociophonetics

Although rather less 'glamorous' than forensic phonetics and its applications related to criminal law, sociophonetic research employs the same acoustic tools to identify socially motivated phonetic variation. A wonderful example of sociophonetic research is the considerable body of work done by British sociophonetician Jonathan Harrington (with Australian and New Zealand colleagues) on changes to the speech of Queen Elizabeth II over several decades from the 1950s onwards, using her annual Christmas broadcasts for data. In the first study, the Queen's monophthongs from the 1950s to the 1980s were compared with data from a small corpus of four female speakers of Standard Southern British English from the 1980s (held to be a more modern form of the ‘prestigious’ Received Pronunciation accent). Formant frequency analysis indicated that the Queen’s vowel space had expanded vertically (that is, into a greater F1 range), and to a lesser degree compressed horizontally (F2), which was seen to be in keeping with a more mainstream form of RP. A later study on her diphthongs showed that these were shifting in a similar direction. Thus, even the Queen – the archetypal speaker of the Queen’s English – is able to modernise her pronunciation. Harrington is quoted in *The Guardian* (21 December 2000) as saying, ‘We conclude that the Queen no longer speaks the Queen’s English of the 1950s.’

Sound change

Acoustic analysis of formants is also used to identify changes to vowel pronunciation in much larger speaker groups within a population. Until a change becomes established, it is simply called *variation*, but once it is on the way to becoming permanent, it is referred to as *sound change*. Interestingly for two nations whose European origins were very similar, New Zealand and Australian English vowels have started to travel in opposite directions in the last couple of decades. Studies have shown that the short front vowels of Australian English are now lowering (that is, they are more open), so for young Australians *dress* now sounds more like *drass*, and *had* more like *hud*. But, in New Zealand English, the short front vowels have shifted upwards and become closer, and the vowel in *hit* has centralised to the point of becoming ‘schwa-like’ (as in *fish and chips* /fəʃəntʃəps/). To the Australian ear, it seems as though these New Zealand vowels are upside down – the perception is that over there you’re invited to a *widdeng* not a *wedding*. Add to that the ‘*near-square* merger’ in New Zealand English and you might go to the *heredrisser* before *hiddeng* to the *widdeng*. Of course, New Zealanders have a similar twist on the Australian accent! For example, the word *wedding* when pronounced by a speaker of Australian English can come across as either *weeding* or *wading*. And the number *six* is notable because both groups report hearing the other’s pronunciation as *sex*. These examples of sound change were first observed over quite a long period as auditory impressions, and subsequently verified by recording speakers under laboratory conditions and measuring the formant frequency values from the spectrograms.

Koalas, kiwis and cognitive bias

There is a lot of research that has been carried out on the effect social information has on our perception of sounds. Studies show, for example, that if listeners believe speakers to have certain characteristics, they will expect them to sound a certain way. This is similar to the priming effects we described earlier (recall that in the case of incoherent recordings, context and our own knowledge help our brains to make sense of the noise we hear). Research shows that even objects can prime listeners to hear speakers in a certain way. In a brilliant perception experiment, Hay and Drager surreptitiously exposed their participants to different stuffed toys – either kangaroos and koalas (from Australia) or kiwis (from New Zealand). Even though these participants were completely unaware that the toys had anything to do with the experiment, their perceptions of vowel tokens as being more ‘Australian’ or more ‘New Zealand’ were influenced by the choice of toy.

Figure 7.22 The Australian koala and New Zealand kiwi

In research on language change, it is the phonetician’s role to track the detail. Acoustic analysis lets us pin down suspected modifications as they are actually happening, including the so-called /el/–/æl/ merger currently taking place in Australian English. Again, this work involves Debbie Loakes and colleagues from The University of Melbourne, and they describe this phenomenon as the lowering of the /e/ vowel before laterals, resulting in the disappearance of the contrast between /e/ and /æ/ in this context. This means that the words in pairs such as *celery–salary* and *Ellen–Alan* become homophones (they come to sound the same). This collapsing of the /el/–/æl/ distinction is often identified as characteristic of speakers from Victoria, especially Melbourne, but it has also been identified in the speech of New Zealanders.

So far, Loakes and her co-researchers have found that in the south of Victoria people are more likely to perceive and produce /el/ vowels with lower quality than someone in the north. The vowel chart from their data (Figure 7.23) shows a 31-year-old female speaker of Australian English from Melbourne saying *hit, het, hat, hut, hot, hoot* (nonsense word rhyming with *put*), and six repetitions each of *hell* and *Hal* (in the ellipses). As you can see, the /el/ and /æl/ tokens overlap, demonstrating that there is little difference in the pronunciation of these words, and crucially that /e/ and /æ/ do contrast in non-lateral environments such as *het* versus *hat*. Just as interestingly, in the prelateral environment (that is, before /l/), the /el/ tokens are produced in a lower and more retracted position than /e/ in the word *het*. If you go to the website accompanying this textbook you can hear a couple of this speaker's sample words for *hell*.

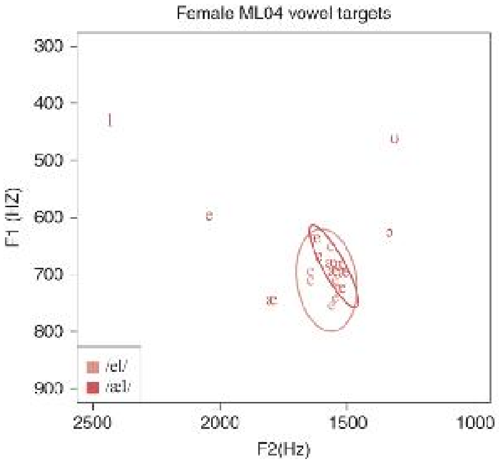


Figure 7.23 A Melbourne speaker's pronunciation of *hit, het, hat, hut, hot, hoot*

Collapsing vowels in Malbourne

Aidan Wilson (former editor of Fully Sic) posted the following comment on the collapse of /el/ and /æl/ (https://blogs.crikey.com.au/fullysic/2010/11/23/walcome-to-malbourne/):

I was out at dinner with a group of people, one of whom happens to have been called Ellen (who is not from Melbourne). She remarked that since she moved here, she's noticed a lot of people calling her Allan. To which someone else (who was from Melbourne) responded that she couldn't even hear the difference between the two. Being the constant documentary linguist that I am, I asked her for a couple more words.

What's the green stalk-like vegetable that people put peanut butter on? I asked.
Salary she replied.
Where do evil people go when they die? I asked.
Hal she answered.

Further reading

Good introductory works on phonetics include Catford (2001), and Ladefoged and Johnson (2011). Ladefoged and Disner (2012) provide a more detailed account, and Ladefoged and Maddieson (1996) is an excellent resource for exploring the full diversity of human speech sounds. The *Handbook of the International Phonetic Association* (International Phonetic Association, 1999) is also a valuable resource. Accessible resources on Antipodean English include Harrington, Cox and Evans (1997), Mannell, Cox and Harrington (2014) on Australian English, and Hay, Maclagan and Gordon (2008) for New Zealand English. See also the summary statements in Cox and Palethorpe (2007), and see Bauer et al. (2007), and Bauer (2015). Crystal (1994) documents elements of syllable-based rhythms in major varieties of English, and for the expansion of this and other Polynesian features in New Zealand English, see Bell and Gibson (2008). More detailed accounts are Cox (2012), and Bauer and Warren (2004). The material here follows these sources. The Australian vowel and diphthong charts follow Cox and Palethorpe (2007), while the New Zealand charts follow Bauer et al. (2007). For more on sound symbolism see Lowrey and Shrum (2007), and Klink (2000), and the story of the Ford Edsel is from Lacey (1987). Benczes (2019) provides a wonderful foray into the important role of phonological motivation in language use (such as phonesthemes and the correlation between 'smallness' and high front vowels). Baldwin and French (1990) describe some well-known cases involving forensic phonetics. See also Olsson (2004), and Coulthard and Johnson (2007). For more on 'suggestibility' and the fall-out of simple priming in criminal trials, see Ridley, Gabbert and La Rooy (2012), and Fraser (2018). At Forensic Phonetics Australia (https://forensicphonetics.com.au), you can find details about general forensic phonetics, as well as explanations and audio demonstrations illustrating how transcriptions of covert recordings can lead to wrongful convictions. Papers on the Queen's Christmas speech include

Harrington, Palethorpe and Watson (2000, 2005). The /el/–/æl/ ‘merger’ is described in Loakes et al. (2014). For the research on priming with stuffed toys, see Hay and Drager (2010).

Exercises and discussion points

1 Changing places

Commence with the English word *clue* each time and change it into another English word as indicated below. Give an IPA symbol and a brief articulatory description for the new sound each time. (Do not change any more than the instructions require.)

(a) Change the voicing of its initial stop.

(b) Change the voicing and place of articulation of its initial stop.

(c) Change its lateral into another approximant.

(d) Change the place of the initial stop and turn it into a fricative.

(e) Change its vowel into a diphthong.

2 Translation

Translate these phonemic transcriptions (in Australian English) back into their written forms.

(a) /ɪgzæmz/

(b) /hɔlədæɪz/

(c) /jʉːnəvɜːsətiː/

(d) /teksts/

(e) /kwɔləfəkæɪʃən/

(f) /lɪsən/

(g) /fəʉniːm/

(h) /kæmpəs/

(i) /ɹəʉtæɪt/

(j) /skɑept/

3 Phonemic transcription

Transcribe the following words based on your own accent.

- **(a)** peak
- **(b)** catch
- **(c)** luck
- **(d)** rude
- **(e)** round
- **(f)** speedy
- **(g)** skilled
- **(h)** bathers
- **(i)** frown
- **(j)** physical

4 Transcription practice

Make a phonemic transcription of the segment of speech we have set aside for you in the online resources. Compare it with the transcription also provided. What sort of information is present in your phonemic transcription that is lacking in the transcription into the English orthography? Do both versions match in terms of word boundaries?

5 Swear words

Imagine hitting your thumb with a hammer or being cut off in traffic. Which would be a more effective emotional release *Rowng!* or *Grock!*? Here's an activity to help you explore the phonetics of swear words. Make a list of all the swear words you know, then answer the following questions.

- **(a)** What consonant sounds appear in your list?
- **(b)** How many of each type?
- **(c)** What types of distribution patterns do you see? (Are there more examples of certain sounds in particular places?)
- **(d)** What do they have in common – for example, are there more stops or approximants?
- **(e)** Now take a look at the vowels. Which feature most frequently?

(f) Why do you think these sounds (both consonants and vowels) would be an advantage for swear words?

Research project

Tips and resources

The phonetics of drunkenness

Read the article by Barbara Heigl entitled ‘[s] Under the Influence of Alcohol’, published in *Working Papers in Linguistics*, vol. 8, Fall 2011. Then view the video clip of a speech given by John Wayne in an apparently intoxicated state.

In what ways is his speech affected? Draw on the information in section 2 of Heigl’s paper when conducting your analysis, and write a report presenting your findings.

For more activities, please go to www.cambridge.edu.au/academic/linguistics.

8

Phonology: The sound system

◆

You can't be happy with a woman who pronounces both d's in Wednesday.

Peter de Vries (1981), *Sauce for the Goose*

8.0 Introduction

Sounds can be examined from two perspectives: the **phonetic** and the **phonological**. Phonetics, as you've witnessed in the preceding chapter, is the study of the way we make, transmit and receive speech sounds. It is concerned with describing the physical units of sound produced by the speech organs. Such sound units are called **phones**. In contrast, phonology is the study of the sound systems and patterns of sounds that occur in languages. It seeks to determine the distinctive sound units in a language by finding out which of these produce differences of meaning within that language. When people listen to utterances, they are concerned with the significant aspects of the sounds, not the total awareness of how each sound is made. What they are really focusing on is how the distinctive sounds in particular languages convey differences in meaning. These meaningful sound units are called **phonemes**. In this chapter, we shift our attention away from the physical aspects of phonetic descriptions to this more psychological basis of the sound system, or **phonology**, of a language – here, we're more interested in what's going on in people's heads than in their mouths.

Recall that to help separate phonetic and phonological approaches, linguists use different conventions to represent sounds. Square brackets [] signify that sounds are being discussed from a phonetic point of view (that is, as phones or physical speech sounds), and slashes / / indicate that they are being discussed from a phonological point of view (that is, as phonemes, or the parts of the sound system of a language that speakers carry around in their heads). One final useful set of brackets is the angle brackets < >, which show that sequences are written using the normal writing system rather than the International Phonetic Alphabet (IPA). The three versions of the word *stand* would be [stæ̃nd] (the squiggle shows the nasal has drifted back over the vowel), /stænd/ (omitting the nasality that is simply a by-product of the neighbouring [n]), and <stand> in normal English spelling.

8.1 Phonemes, phones – and allophones

Let's begin with a concrete example. Consider the sound /t/ as it appears in the words *ton, stun, bet* and *better*. In each of these words, the consonant has quite different pronunciations, depending on its position. At the beginning of *ton*, it is followed with a little puff of air, known as **aspiration** (if you lit a match and said the word *ton* you'd likely blow it out). After /s/ in a word like *stun*, this puff of air is suppressed (you'd be uphill blowing the match out here). At the end of a word like *bet*, /t/ is almost swallowed; this is an unreleased stop (the organs of speech aren't opened again after the sound is produced). In the middle of a word like *better*, /t/ is pronounced like a hastily articulated /d/; this is known as **tapping** or **flapping** (the tip of the tongue rapidly strikes the alveolar ridge). These 't's (phones) are all very different sounds, and a phonetic system (such as the IPA) represents each of them with different symbols: aspirated [tʰ], unaspirated [t], unreleased [t̚] and flap [ɾ], as shown in example 8.1.

(8.1)

ton [tʰɐn]

stun [stɐn]

bet [bet̚]

better [beɾə]

As interesting as these fine phonetic details are, we don't need this sort of information in our writing because it doesn't make any difference to the meaning or our understanding of the word. This variation is not significant – if a speaker didn't aspirate the /t/ at the start of *ton*, it might just come across as an odd way of saying the word. When speakers think of the 't' sounds in the words *ton, stun, bet* and *better*, they think of them as the same sound; their brains filter out the phonetic details and 'hear' only the phoneme, which we represent here as /t/.

The different phones, or phonetic shapes of this phoneme, are predictable. Speakers make fine adjustments (such as aspiration) automatically as part of their language ability, allowing different pronunciations of the phones to be recognised simply as versions the same phoneme /t/. The technical term for these different versions is **allophone** (the term comes from the Greek: *allos* 'other' + *phone* 'sound'). Thus, in example 8.1, [tʰ], [t], [t̚] and [ɾ] are allophones of the phoneme /t/. Compare this to how you perceive colour – you may not notice, unless your attention is drawn to it, that the carpet you think of as uniformly blue has all sorts of blues in it, depending on different conditions of light and shade.

The initial consonants of the words *ton, done, bun* and *shun* are distinctive sounds in English, and thus phonemes (/t/, /d/, /b/ and /ʃ/). Switching these phonemes results in saying a different word: *ton* /tɐn/, *done* / dɐn/, *bun* /bɐn/, *shun* /ʃɐn/ and so on. This is what we want a spelling system to capture.

In Chapter 7, we introduced two different types of IPA transcription: a **narrow transcription** provides all sorts of phonetic details (aspirated, unaspirated, flapped, unreleased, and so on), while **broad transcription** indicates only the phonemes (just the amount of detail required to distinguish between words). Take the phrase 'phonetic transcription' and consider how it might look in these two scripts. There will be some symbols here that you won't recognise, but you will be introduced properly later in the chapter.

Broad (phonemic)	/fəˈnɛtɪk trænˈskɹɪpʃən/ (abstract – as occurs in someone's head)
Narrow (phonetic)	[fəˈnɛɾɪk tʰɹæ̃nˈskɹ̥ɪpʃn̩] (concrete – as actually pronounced)

Variation can be stylistic (e.g. is it slow, careful speech or faster, more casual speech) and social (e.g. what is the socio-economic background of the speaker), as well as linguistically conditioned (e.g. what are the neighbouring sounds). So far we have seen examples of linguistically conditioned variation. Linguistic conditioning refers to the context – for example, the position of /t/ in a word and the types of sounds in its vicinity. This type of conditioning is probably the hardest to spot, at least for native speakers, because they are not usually conscious of these differences. Sounds may also be pronounced differently depending on the regional variety of English that is being spoken. For many New Zealand dialect speakers, *cheer* and *chair* are pronounced the same (that is, they are homophones). For some Americans, *father* and *bother* rhyme, and *cot* and *caught* are homophones. In Atlanta, *boil* and *sky* sound more like [bal] and [ska]. Clearly, none of this variation could (or should) be captured in a spelling system! For a start, how would you reconcile the different needs of the R-ful (rhotic) and R-less (non-rhotic) dialects we mentioned in Chapter 7? Writing words how they sound means that all those dialect specific homophones will need to be negotiated (for example, *father* and *farther* in Australian and New Zealand English).

English spelling – is it all that bad?

For many years, people have been drawing attention to the inconsistencies of the English alphabetic writing system, often in rather peculiar (and misleading) ways. In the 1930s, R.E. Zachrisson, Professor of English at the University of Uppsala, reported that on the basis of all potential combinations of sounds, it was possible to spell the word *scissors* 596 580 different ways. Another person has worked out that *foolish* could be spelt 613 975 different ways. Irish playwright George Bernard Shaw once famously showed that *fish* could be written as *ghoti* (and we'll leave you to figure out the basis on which he made this calculation). Someone else once even determined that 50 million schoolchildren spend 10 million hours daily on learning the English spelling system. (If that is not enough, this person went on to calculate that this was roughly equivalent to the number of inches between London and Hull!)

Sure, English spelling is not optimal – we've already established that: twelve different ways of representing /t/ is hardly ideal. But, the system is not as bad as people like to make out. In fact, more than 80 per cent of words are spelt according to regular patterns, so wholesale change is not what we want. Modest adjustments could, however, be made without any major upheaval. We could promote more regular spellings wherever they are already familiar as alternative spellings. To introduce the *-or* spelling across the board would be a painless reform, especially since it only affects about 30 words. A few straightforward reforms like these would help tidy up the system considerably. But, it is doubtful whether reforms (as sensible as they might be) such as *tung* or *bred* will ever catch on. Whenever spelling reform boards start pushing for these sorts of reforms, reactions are immediately hostile. It is difficult to imagine that psychologists would ever be happy with *sikolojy*. You can only go so far with simplified spellings!

To sum up the chapter so far, phonetics gives us a way to describe the phones of languages and how they differ, so it deals with unprocessed, raw sounds. Phonology looks at the structure of sounds and how they function in the system associated with a particular language or dialect; it identifies the phonemes (or distinctive sounds) of a language and how the phones pattern as allophones of these phonemes. We will return to the ways in which we identify what are and are not the distinctive sounds in a language a little later, and we will also examine some of the processes that lie behind the variation. For the moment, consider some of the things you can do with this sort of information.

Why bother with allophones and phonemes?

Throughout this book, we have been highlighting some of the practical applications of linguistics, and at this point you might be wondering about the reasons for studying phonology (apart from the intrinsic interest and sheer delight, of course!). Chapter 7 introduced you to some of the real-life uses of phonetics – most obviously in the world of forensic linguistics – and you will perhaps be comforted to know that phonology is also important to a number of applied subfields of linguistics.

There are a number of jobs that should not be attempted without a firm understanding of how phonological systems work. Bear in mind that the majority of the world's languages still do not exist in written form, so many linguists are involved in devising orthographies for these languages. Another area where phonology is important is communication disorders. Speech therapists work with people whose language ability is somehow different from the expected (perhaps due to developmental problems or to illness that damages language skills that previously had been fully developed), and they develop techniques that can help these people to improve their communicative abilities. This work can involve many aspects of phonology, including voice disorders, problems of intonation and difficulties around the articulation of sounds and their distribution.

Probably the most obvious practical advantage to knowing about phonemes and allophones, though, is in language teaching, where a thorough understanding of phonology is useful to both teacher and student. Phonological interference – or what in general terms is called ‘a foreign accent’ – involves the transferral of established speech habits into a second language. This is understandable: we naturally hear and perceive sounds and sound patterns through the filter of our own language.

The Russian phoneme /r/ is pronounced as a trill (with the tip of the tongue vibrating against the teeth ridge), whereas in varieties of English in Australia and New Zealand /ɹ/ is a postalveolar continuant (with the curled tongue tip approaching the area behind the teeth ridge). The English pronunciation tends to carry over into Russian for English speakers (and vice versa). In other words, interference can involve the physical aspects of sounds production, and it can be very helpful for learners to be given articulatory explanations of what is involved. But, this is only part of the story.

Interference also involves trickier aspects of the sound system. English has voiced and voiceless (inter-) dental fricatives ([θ] and [ð]), but these sounds aren’t found in many languages (suggesting there is something intrinsically difficult about them). Russian speakers learning English substitute [t] and [d]; German speakers substitute [s] and [z]. The interesting aspect of this is the nature of the replacement sounds: German speakers could opt for the stops, but they maintain the fricative feature (so the mode of articulation); Russian speakers could go the fricative route too, but they opt to maintain more of the English place of articulation (you’d have to look closely at the sound systems of these two languages to know why speakers make different choices here).

Phonemic distinctions of the first language are often imposed on the second language, and this contributes strongly to a perceived foreign accent. Language X might lack a distinction that is present in language Y, and if language X speakers impose their sound patterns on language Y they will under-differentiate sounds. Korean, for example, has one liquid consonant, which can be realised as either a lateral [l] or a flap [ɾ]; English has two liquid phonemes, /l/ and /ɹ/. It requires extra effort for speakers to perceive and produce sounds accurately if they lack the contrast in their first language.

Restrictions on the arrangements or sequences of speech sounds can also prove to be stumbling blocks for learners of foreign languages who might fail to acquire the patterns of their new language or, more usually, end up transferring the rules of their primary language into the new language. Take something as straightforward as word-final position. In Chapters 5 and 11, we describe how German speakers devoice final stops and fricatives; many carry this rule over to English (so *trend* is pronounced as [trent]). English has three nasal phonemes – /m/, /n/, /ŋ/ – that can all appear word-finally. Spanish has restrictions on final nasals and Spanish speakers can have trouble identifying and pronouncing nasals other than /n/ in final position (so *long* might be pronounced as [lɔn], not [lɔŋ]). Japanese is much more restricted than English regarding what consonants can and cannot occur at the end of a word or syllable, preferring words that end in vowels. In English, /h/ is the only consonant that can't appear word-finally. The only final consonant that doesn't present difficulties for a Japanese-speaking learner of English is final [ŋ] (as in *bling*); everywhere else, the natural solution is to insert an extra vowel at the end (thus *net* [net] might become [netto]).

Phonology and the law: Trademark infringement

Phonology is a big part of forensic linguistics, particularly in branding. Trademarks are words, symbols or other devices used to identify goods or services and distinguish them from others. So, they have to be distinctive if they are to prevent consumer confusion.

Names that are generic or descriptive are the hardest to protect against use by other companies – for example, the chewing gum company Wrigley's wasn't able to trademark the word *double-mint* precisely because its dictionary meaning was the meaning used for the product. The strongest brands, therefore, involve more fanciful creations because it can be argued that they are uniquely identifiable with the product. 'Suggestive' names involve common nouns, but they allude to a quality or characteristic of the product in some way: they require imagination – so we have *Playboy* (magazine), Shell (gasoline) and Greyhound (bus company). 'Arbitrary' names (easier to protect) are common usage words with meanings unrelated to the product – for example, Apple (computers) and Ivory (soap).

It sometimes happens that one company feels that the trade name of another company (with a similar product) sounds too much like its own and there is potential for confusion. A famous 2002 case was Victoria's Secret (American lingerie giant) versus Victor's Little Secret (a small sex shop in Kentucky owned by Victor Moseley). Even though Victor Moseley changed the name of his shop from 'Victor's Secret' to 'Victor's Little Secret', Victoria's Secret successfully claimed that the linguistic similarity was too great, and this was blurring and tarnishing its brand.

There are many examples where phonology is used to establish the strength of names and the likelihood of confusion. It is easy to show that 'Lexus' and 'Lexis' are pronounced identically in normal speech, because the vowels in the second syllable are unstressed, and therefore are both realised as schwa [ə]; or that the name 'Lectronic' (for electronic goods) can't be protected because it is a usual pronunciation for the common word *electronic* (the first vowel-only syllable [e] is frequently dropped). Linguists are often called on to analyse the phonological similarity of names. One of us was involved in such a dispute a number of years ago, involving a pharmaceutical product (pain medication); the company that produced the established product argued that the proposed trademark was too close to its own. From their appearance, the trade names were visually very distinct – they were spelt very differently and they were even different lengths. However, it turned out that the names were almost identical in pronunciation because of the transitional effects of normal speech. All sorts of things happen to sounds in rapid speech.

8.2 Sounds in connected speech: Phonological processes

Chapter 7 showed how sounds change when they are part of normal discourse. When we speak, we don't just speak in single words, but in groups of words, which are continuous, and they may or may not have pauses between them. Because of this, a number of things can happen to alter the shape of words, and in this section we consider in detail some of the most common processes that affect sounds in connected speech.

Changes that alter syllable structure: Deletion, addition and reordering of sounds

Deletion occurs when sounds are left out, as frequently happens in everyday casual speech. Unstressed vowels may be dropped (as in *ev(e)ry, t(e)rrific* and *choc(o)late*), and consonants might also disappear, especially if they are made at the same place of articulation as their neighbours: *landlord* is generally pronounced as [lænloːd] – the /n/ and the deleted /d/ are alveolars; similarly, we have

Wednesday as [wenzdæɪ]). French has a regular rule of final consonant deletion where the following word begins in a consonant; for example, [pətit ananas] ‘small pineapple’ versus [pəti garsɔ̃] ‘small boy’. Māori also shows regular consonant deletion in verb roots. Consider the forms in Table 8.1.

Table 8.1 Māori passive formation

Root	Translation	Passive	Translation
mau	‘to carry’	*mauria*	‘be carried’
awhi	‘to embrace’	*awhitia*	‘be embraced’
hopu	‘to catch’	*hopukia*	‘be caught’
kimi	‘to seek’	*kimihia*	‘be sought’
inu	‘to drink’	*inumia*	‘be drunk’
tohu	‘to point out’	*tohungia*	‘be pointed out’

Originally, the active verb forms were *maur, awhit, hopuk*, and so on, but when speakers stopped pronouncing the final consonants, things got complicated. For a while there would have been competing forms (allomorphs; see section 5.4) with or without final consonants (for example, *maur–mau, awhit–awhi* and *hopuk–hopu*) before the system settled down. The passive suffix was initially *-ia*, but when consonants in the final position dropped off the stem (so *maur* became *mau* and *awhit* became *awhi*, and so on), they attached instead to the suffixes (*-ria, -tia, -kia* and *-hia*). Since there is no obvious motive for these differences, they appear to be similar to English irregularities in plural formation, such as *ox–oxen* and *foot–feet*. This is how Māori speakers view this aspect of their language. As a result of shifting the morpheme boundary, speakers now have to learn a whole bunch of different passive endings (whereas once it was simply *-ia*). These irregularities in Māori are the result of a regular process of deletion.

Not just segments but entire syllables can be deleted; this is a process known as **haplology** (or ‘haplogy’!). This is a regular rule of English adjectives when the suffix *-ly* is added to a stem that already ends in *-le*; for example, *gentle* + *ly* becomes *gently* not *gentlely*. Typically, the deleted syllable is identical or similar to a neighbouring syllable, as is the case with *library*, which can be heard pronounced with three syllables, [lɑebɹəɹi], or with two, [lɑebɹi].

Addition of sounds can happen in various ways. Sounds may be added as the various parts of the vocal tract move from one sound to another; this is known as insertion or **epenthesis**. In the case of consonant addition, most examples are the result of some kind of mistiming. Say the words *prince* and *prints*. You probably think you pronounce them differently, but in fact you don’t. It’s just the spelling that suggests this. For both, you would say something like /pɹɪnts/. Whenever we have a cluster of a nasal and a following consonant, as in the word *prince* /pɹins/, we usually insert an oral stop that is produced at the same place of articulation as the nasal (in this case /t/). It's difficult not to do so.

As Figures 8.1 and 8.2 show, our soft palate is lowered for /n/ and the air passes through the nose. Because the following consonant /s/ is oral not nasal, we need to raise the soft palate to stop the air escaping through the nose. But if the soft palate is raised too soon – in other words, before the tongue position has changed – an oral stop is produced in the same place of articulation as the nasal stop.

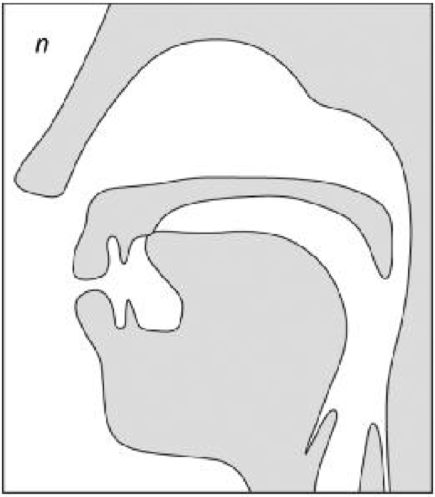


Figure 8.1 Articulation of an alveolar nasal: soft palate is lowered; air passes out through the nose.

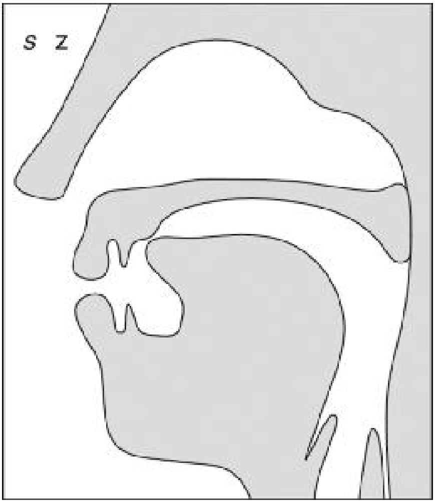


Figure 8.2 Articulation of an alveolar fricative: soft palate is raised; air passes out through the mouth.

As another example, take a non-standard pronunciation, [fæmbli] for *family*. To articulate the nasal [m], the soft palate is again lowered (so the air can escape through the nose). In anticipation of the [l] in *family*, you might raise the soft palate too soon, in which case you will automatically insert a little [b] and the word comes out as *fambly*. Examples like *fambly* (also *chimbly*) are associated with child language, but the [b] in *tremble* and *humble* was added in precisely the same way and stuck (compare *tremulous* and *humility* with no extra consonant).

The addition of vowels has quite a different motivation. There's a lot of evidence to suggest that speakers do actually prefer a consonant–vowel syllable structure and, in keeping with this, we often find vowels being inserted between consonants, usually to help the pronunciation of what are otherwise difficult combinations. It's a kind of consonant cluster simplification. You've probably encountered non-standard pronunciations, such as *nuclear* as *nucular* (on the pattern of other -*cular* words, such as *binocular* and *particular*). Other examples include non-standard pronunciations such as 'fillum' [fɪləm] for *film*, 'athelete' [æθəlit] for *athlete* and 'umbarella' [ɐmbəɹelə] for *umbrella*. Liquids (the usual suspects) are causing the difficulty here, and another option is to drop them. Many speakers now pronounce *film* as [fɪʉm] (compare standard pronunciations of *walk* and *calm*, which have no hint of their original [l]).

Mr Rabbit

Speakers of 'R-ful' dialects like American English are often puzzled that Australians and New Zealanders appear to insert /ɹ/ in the middle of the phrase *law and order*, so that it sounds more like 'law-r-and-order' [loːɹənoːdə].

As noted in Chapter 7, speakers of 'R-less' (or non-rhotic) dialects, such as Australian and New Zealand English, only articulate 'r sounds' that occur before a vowel. So /ɹ/ at the end of words like *mister* isn't usually pronounced. However, if followed by a word beginning with a vowel, /ɹ/ reappears – for example, *Mr and Mrs* is pronounced 'mister-r-and-missus' [mɪstəɹənmɪsəs]. This is called a 'linking r', and it is the inspiration of the so-called intrusive 'r' in *law and order*. Words such as *mister, soda* and *idea* end with the same schwa vowel [ə], so if you say 'mister-r-and-missus' with /ɹ/, why not 'soda-r-and-ice' and 'idea-r-of it'? So /ɹ/ sneaks in this way. Speakers of Australian and New Zealand English have now extended the ‘linking r’ rule to cover any situation where two vowels come together – sometimes even inside words (so *drawing* might be pronounced ‘draw-r-ing’ [dɹoːɹɪŋ]).

It seems unfair that poor old ‘intrusive r’ should get such bad press, when the language is full of interlopers that ensure smooth transitions from one sound to another. No one notices the inserted glides in words such as *going* (‘go-w-ing’) and *seeing* (‘see-y-ing’). And it was certainly unfair when, in 2010, the Australian public accused then Prime Minister Julia Gillard of deliberately slurring her speech so she could call Mr Abbott (the opposition leader at the time) Mr Rabbit. After all, there is no /ɹ/ intruder in Mr Abbott. This is the perfectly standard ‘linking r’ – Mr-r-Abbott is no different from ‘mister-r-and-missus’.

Metathesis (or ‘methatesis’) is based on the Greek word for ‘transposition’, and it involves a reordering of adjacent segments within words. It starts off as a kind of mistiming, usually involving awkward or difficult sequences of sounds, such as the notoriously unstable sonorants (nasals and liquids). English speakers can occasionally be heard to say *renumeration* for *remuneration*; Japanese speakers say /fuiNki/ for /fuNiki/ ‘ambience, mood’; German has both the metathesised and original forms of the proper name *Birgit/Brigitte* (compare English *Bridget*); while Māori metathesis usually involves nasals, such as *rango/ngaro* ‘fly’. Clusters of [s] and plosive consonants, such as [p] or [k], are also typical shifters for English. You’ve probably heard small children say *pasghetti or bisghetti* for *spaghetti* and *hostipal* for *hospital*. These can remain childhood errors, or one-off slips of the tongue, but when a number of speakers keep making the same sorts of adjustments, tongue slips can endure.

Metathesis is typically sporadic and not productive; however, occasionally the addition of morphemes can trigger a regular reordering of the segments. Faroese (a North Germanic language spoken on the Faroe Islands and in parts of Denmark) metathesises clusters /sk/when the neuter suffix *-t* is added (producing a sequence of consonants that is more user-friendly).

Table 8.2 Metathesis in Faroese

Masculine singular	**Translation**
/fɛsk-ʊr/	‘fresh’
/rask-ʊr/	‘energetic’
/frɛŋsk-ʊr/	‘French’
/daŋsk-ʊr/	‘Danish’
/baisk-ʊr/	‘bitter’

Processes that involve the modification of sounds

Perhaps the most obvious aspect of connected speech is that sounds often change their shape depending on the company they keep. By far the most likely way that they will change is to become more like neighbouring sounds. This is known as **assimilation**. Recall that in Chapter 7 we described consonants according to three parameters: place of articulation, manner of articulation and voicing (sounds as either voiced or voiceless). When sounds assimilate, they change according to one or more of these parameters. Nasals are forever altering their appearance to blend into their environment (they are the chameleons of the phonological world!). For example, in normal rapid speech, most speakers pronounce *pancake* as [pæ**ŋ**kæɪk], assimilating the [n] to the following velar [k]. In 'pangcake' the /n/ is anticipating the place of articulation of the velar /k/, becoming velar /ŋ/. These features are shown in example 8.2.

(8.2)

[n]	[ŋ]	[k]
voiced	voiced	voiceless
alveolar	velar	velar
nasal	nasal	stop

This type of assimilation is called **anticipatory assimilation**. English nasals always anticipate the place of articulation of the following sound and convert to it – in other words, they become homorganic (a fancy way of saying 'sharing aspects of articulation') with that sound. Clear examples can be found in negative prefixes such as *in-* and *un-* (and compare the quality of the nasals in the *com-* prefix, as in *compel* and *contend*). More instances are shown in example 8.3.

(8.3)

[ɪ**n**edəbəl] *inedible*

[ɪ**nd**isənt] *indecent*

[ɪ**nt**ɔləɹənt] *intolerant*

[ɪ**mp**ɔsəbəl] *impossible*

[ɪ**mb**æləns] *imbalance*

[ɪ**ŋk**ɹedəbəl] *incredible*

[ɪ**ŋg**loːɹiəs] *inglorious*

The last two forms show the nasal becoming velar before velar consonants, but this isn't reflected in the spelling (/ŋ/ is a recent phoneme, and we don't have a symbol for it). In *illegal* and *irregular*, the assimilation is total: the nasals have given in completely to the following liquids; once they would have been pronounced as a **geminate** (long or doubled) consonant. This kind of consonant assimilation even works across word boundaries: in *handbag,* /d/ is deleted and the word becomes *hambag* /hæmbæg/. We can represent this kind of anticipatory assimilation in a formula: AB → BB.

Assimilation can also work in the opposite direction. In other words, a sound can take on one or more features of the preceding sound. This is called **lag assimilation** and the formula would be AB → AA. This process is less common, but plenty of examples exist in English, especially in rapid speech. Consider the pronunciations of *seven* and *eleven* as */sevm/* and */əlevm/* (where the nasal /n/ has taken on the labial qualities of the /v/). In fact, sounds often borrow from each other in normal rapid speech, so assimilation can work in both directions simultaneously – for example, *seven* and *eleven* might be pronounced as /sebm/ and /əlebm/. (The labial quality of /v/ is transferred to the following nasal, which then gives its stop and bilabial qualities to the /v/, converting it to a /b/.)

Another type of assimilation commonly heard in Australian and New Zealand accents is called **voicing assimilation**. This involves the voicing of voiceless consonants when they occur between vowels in words such as *better* and *atom*, or in a sequence like *get it*. Think about it – your vocal folds are vibrating for the vowel sounds, and it is much easier not to shut off the voicing for the intervening consonant – hence we pronounce these as if they were spelt with a <d>: *bedder, adom* and *gedit*. What we are pronouncing here is really a very fast / d/; as described earlier, this is the alveolar flap represented by a symbol that looks a bit like a fish-hook ([ɾ]). (It is actually the pronunciation of the Japanese rhotic in words such as *hara, akira* and *tora.*) Mali shows a similar process. For example, the word for hand, *tik*, gains an initial voiced stop when a prefix ending in a nasal (in this case, the second person dual-possessive prefix) is added: *ain-dik* 'you two, your hands'.

A final type of assimilation is one that is so common that it has its own label – **palatalisation**. This sort of change affects consonants appearing before front vowels such as /i/ and /e/, and also the palatal approximant /j/ (which you may remember is produced by raising the tongue to the palate directly and so is very like the high front vowel /i/). The raised-tongue position of the vowel or approximant gets transferred onto the preceding consonant. Take a word sequence such as *can't you* [kɐːn**t** **j**ʉː] or *did you* [dɪ**d** **j**ʉː]. The vocal tract is constricted for the articulation of [j], and the consonant [t] or [d] shifts into the same area and produces the palatal affricate [tʃ] or [dʒ], giving the pronunciations [kɐːn**t**ʃə] and [dɪ**d**ʒə], respectively. It is as if a palatal articulation is superimposed on the consonant.

Russian has spectacular examples of palatalisation; as Table 8.3 shows, certain consonants become palatal when they occur before front vowels (the palatal sounds are bolded here, and the palatal quality is indicated by the raised j).

Table 8.3 Palatalisation in Russian

Basic forms	Palatalised forms	Translation
[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]

It can happen that non-neighbouring sounds assimilate, although this is much rarer. A speaker might intend to say *discussing shortly* but come out with *discushing*. Assimilation more usually involves vowels and is described as **vowel harmony** (or **umlauting** in Germanic languages). Mali shows vowel harmony in clitic groups (sets of clitics that bundle together to form words). For example, in connected speech, the conjunction *da* 'and' can take a number of different forms depending on what follows: [də ŋət] 'and they', [du ʊt] 'and we' or [di ŋia] 'and you'.

In German, the assimilation works in the same direction (that is, it is anticipatory); back vowels are fronted before suffixes containing a high front vowel. English has vestiges of this harmony rule in irregular plurals such as *man–men* and *tooth–teeth* (we examine this in more detail in Chapter 11). In German, the process is so significant that it has its own marker (or **diacritic**): two dots above the vowel (known as an **umlaut**). So vowels written as <a>, <o> and /u/ become <ä>, <ö> and <ü> (to indicate fronting). In Table 8.4, we have given both the spelling and the phonetics so the changes are more obvious.

Table 8.4 Umlauting in German

Basic forms	Umlauted forms	Translation

[illegible]
[illegible]
[illegible]
[illegible]
[illegible]
[illegible]

In modern German, the umlaut pattern is much more entrenched than in English and continues to generate new plural forms. However, it is irregular. Not all high front vowels produce umlauted vowels in this way (so *Amt* ‘office’ → *amtlich* ‘official’, not **ämtlich*), and the trigger vowels are no longer restricted to high front vowels (for example, the mid vowel in *-er*, the plural marker in *Mann* ‘man’ → *Männer* ‘men’).

Dissimilation is really the opposite of assimilation: it describes how sounds change to become less like other sounds in the neighbourhood. Most linguists attribute dissimilation to the fact that we find repetition of the same muscular activity difficult. Repeating a sound in a tongue twister (like *She sells seashells on the seashore*) can trigger the speaker to suddenly do something unexpected, creating a different sound.

Dissimilation is behind the non-standard pronunciation [eksetəɹə] for *etcetera* [etsetəɹə] (though there is added reinforcement coming from forms beginning with *ex-* [eks]). Nasals and liquids ([l] and [ɹ]) are particularly prominent in this sort of dissimilation process. For example, some English speakers dissimilate the second nasal in *chimney* → to [l] ‘chimley’, and the first [ɹ] in *February* to [j] → ‘Febyuary’ (compare Russian *fevral´* [fʲevrˈalʲ] also from Latin *Febrarius*). The Modern German word *Kartoffel* ‘potato’ is interesting because it shows a rare example of distance dissimilation involving stops: *Tartuffeln* (from Italian *tartuffeli*, named after the truffle) changed its initial [t] to [k]. Like metathesis, this change is typically sporadic (successions of nasals won’t regularly prompt dissimilated forms, such as *chimley*).

8.3 Determining phonemes

Given that not all the phones in a language signal differences in meaning, how do we determine just what the significant sounds are? This is exactly the question linguists must ask when they are devising orthographies for unwritten languages. Take the example of nasalised vowels. Speakers of English perceive the vowel sounds in *sub* and *some* as being the same, even though acoustically they are poles apart. If you listen very closely, you may be able to hear the difference: [sɐb] versus [sɐ̃m]. The squiggle above the vowel in *some* shows that it is nasalised. The difference is hard to spot because the nasality is not significant. If someone pronounced *sub* with a nasal vowel ([sɐ̃b]), it probably wouldn't be noticed. English doesn't cash in on nasalisation, although there are plenty of languages that do. In French, for example, what distinguishes *sa* [sa] 'his, her' from *sans* [sã] 'without' is contrast between a nasal and a non-nasal vowel.

So, what method can we use to show, for instance, that there are 44 phonemes in New Zealand and Australian English? The general procedure that linguists use is to look at the distribution of the phones.

Minimal pairs

The first step is to search for pairs of words that differ in only one sound in the same position in each word, such as *thy* and *thigh* or *feet* and *foot* (or [sa] 'his, her' and *sans* [sã] 'without' in French). These small sets of words are known as **minimal pairs**. It is the difference between the two particular sounds in these pairs that alters the meaning, and this is direct evidence that the two sounds represent different phonemes. Thus, *thy* and *thigh* constitute a minimal pair, illustrating that /ð/ and /θ/ are phonemes in English; likewise, *feet* and *foot* show that /iː/ and /ʊ/ are phonemes. Unfortunately, it is not always possible to find minimal pairs, as there are often accidental gaps in the vocabulary where a minimal pair could exist but doesn't actually occur. For example, we are not aware of any minimal pairs for [ʒ] and [dʒ], yet we find that each of these sounds contrasts with other sounds in English, such as in the minimal pairs *measure* [ʒ] and *mesher* [ʃ], and *judge* [dʒ] and *fudge* [f]. Because of this, we can say that we suspect that [ʒ] and [dʒ] are phonemes in English, although we would have to do some more digging to be absolutely sure. (But note, this situation is changing with the influence of borrowings such as French *jus* 'thin gravy' – for some speakers this contrasts with *Jew*.)

It isn't surprising that minimal pairs have been described (by Charles Hockett) as 'the analyst's delight' because they do cut down the work; as soon as you discover a minimal pair, you know you have phonemes. Minimal sets are an even better way of establishing the phonemic inventory of a language, so it is good to try to find contexts (or 'frames') that allow the manipulation of a single sound in this way. Let's say we are trying to establish the distinctive vowels (that is, both monophthongs and diphthongs) of Australian English. The frame [b_t] is a useful one; it catches many sounds.

Table 8.5 English minimal sets

Phonetic transcription of vowel/diphthong	**Keyword(s)**
eə	bear
ɪə	beer
ʊə	boor
iː	beat
ɪ	bit
e	bet
æ	bat
ɜː	Burt
ɑː	Bart
ʌ	but
ɒ	bot(tom)
ʊ	book
ɔː	bought
uː	boot
eɪ	bait
aɪ	bite
ɔɪ	boy, boyter
aʊ	bout
əʊ	boat

We haven't done too badly here – we have minimal pairs for most vowels. Occasionally we had to resort to fairly unusual words, such as *boor* 'a rude fellow' and *boyter* 'a vulture'; the three falling or centring diphthongs are uncommon (they also vary in their pronunciation, and are often pronounced as single vowels). Since there is no longer a word *boit* [boɪt] in Modern English (in very early English, it was a form of the verb *to be*), we had no choice but use a **near-minimal pair** here, such as *boy* or *boyter*. It is easy to show that the phonetic context has nothing to do with what vowel appears in these words – for example, we can find other minimal sets (such as *hood, had, hid, head, hide, heard, hoed, hayed* and *hawed*) to confirm this.

In Table 8.6, consider the sets of minimal pairs that between them justify the six-vowel analysis for Mali.

Table 8.6 Minimal pairs of vowel phonemes in Mali

[illegible]	[illegible]
/biŋ/ 'roll.string.present'	/bɛŋ/ 'climb'
[illegible]	[illegible]
[illegible]	[illegible]
[illegible]	[illegible]
[illegible]	[illegible]

To summarise, minimal and near-minimal pairs establish that sounds contrast or are in **contrastive distribution**; in other words, they are phonemically distinct. Voicing is contrastive for English stops, illustrated by minimal pairs such as *pin* and *bin* (distinguished only by the voicing of the initial sounds). English stops can also be aspirated or unaspirated, but this feature is not contrastive in English. It is in Turung, though: the voiceless unaspirated stops /p/, /t/, /k/, voiceless aspirated stops /pʰ/, /tʰ/, /kʰ/ and voiced stops /b/, /d/, /g/ are in contrastive distribution, and this is demonstrated by minimal pairs such as *lpu* 'snake' /ləpuu³/, *lbu* 'poison' /ləbuu²/ and the near minimal word *lphut* 'knee' /ləpʰut²/.

Complementary distribution

An interesting aspect of allophones of a single phoneme is that they usually avoid sharing the same environment. For example, we may find that they never occur in the same positions in words, or are never surrounded by the same sounds. This is termed **complementary distribution**. When we see this pattern, it provides us with evidence that the sounds are not in a position to contrast linguistically, but are different pronunciations of the same phoneme. To show how to determine complementary distribution, we will start with English, then move onto examples from our case-study languages.

Take a sentence like *Lucy likes salty meals*. There is a clear audible difference between the initial laterals of *Lucy* and *likes* (known as a 'clear [l]') and the laterals in *salty* and *meals* (known as a 'dark [ɫ]').

In each case, a partial closure is made by the blade of the tongue against the alveolar ridge; however, for a clear [l] the body of the tongue then runs down and back, as in the production of a front vowel, whereas for a dark [ɫ], the body of the tongue is higher and shaped like a high back vowel. These 'clear' and 'dark' versions of /l/ are allophones; they are non-significant variants that speakers automatically produce without thought (as we will see later, they are triggered by the context in which they appear).

In fact, there are all sorts of different pronunciations of /l/. Another is the 'syllabic [l̩]' – it takes on the vowel quality and forms a syllable all on its own (consider the lateral in words such as *bushel* and *paddle*). English speakers think of all these different sounds as simply /l/. So, we need to uncover the different phonetic environments that trigger the different hues and tints of this phoneme. Consider the data set in Table 8.7, showing three different versions of /l/.

Table 8.7 English laterals

[illegible]
[illegible]
[illegible]

[illegible]
[illegible]
[illegible]
[illegible]

Take note of where these three sounds appear; consider their position in the word (or syllable), what comes before them and after them, and also any suprasegmental differences (such as stress). In this example, the complementary distribution of allophones is a matter of position in the syllable. The ‘clear l’ always occurs at the beginning of syllables (before stressed vowels, as in *laugh* and *polony*). In contrast, the ‘dark l’ always occurs towards the end of syllables (following stressed vowels, as in *silly* or *milk*). ‘Syllabic l’ only ever occurs word-finally and after a consonant (and is a syllable on its own, as in *puddle*). Since these sounds never occur in the same context, they are in complementary distribution and are therefore allophones of the one phoneme, /l/, which we can represent as shown in Figure 8.3. Here we have added yet another allophone of /l/, a voiceless version, [l̥], that occurs after voiceless stops and fricatives (so words such as *play* and *slope*).

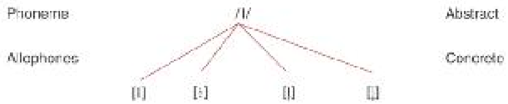


Figure 8.3 The English phoneme /l/

The symbol we choose to put in the // brackets to represent the phoneme is usually the allophone that occurs in the most places (that has the widest distribution), in this case /l/. If that isn’t obvious, then you might decide on the basis of which sound occurs more frequently in the world’s languages. But, it doesn’t really matter what you put inside these brackets because the phoneme is an abstract thing, something you carry around in your head; what are real (in other words, what are pronounced) are the allophones.

Where have all the ‘l’s gone?

Liquids are notoriously unstable sounds. You’ve just seen how /l/ changes depending on where it appears in a word: there are ‘dark’, ‘clear’, ‘syllabic’ and ‘voiceless’ versions. There is yet another variant on the rise, too – a back vowel, [ʉː], which may or may not be rounded or labialised. The contexts that promote this version involve backness (of adjacent consonants and vowels) and syllable position, so you hear it in final clusters (*milk* [mɪʉːk]), end position (*pill* [pɪʉː]) and syllabic environment (*buckle* [bɐkʉː]).

The consonant /l/ has a long history of such disappearing acts. Early on, it disappeared between back vowels and labial or velar consonants. Conveniently, English spelling is conservative and preserves evidence of the earlier pronunciation of words such as *walk, calf* and *calm*. We can date L-drop to around 1500 (from spellings of *walk* as *wauk*), though even into the seventeenth century people were recommending the ‘L-ful’ pronunciations as the correct ones. James Howell (1662) wrote that, ‘*A,* when it comes before *lm* somtimes drowneth the *l* and turneth to an *u,* as Calme is pronounced Caume, Psalme Psaume, Balme Baume, &c.’. In fact, /l/ continues to be drowned by all sorts of sounds, not just by ‘a’ as Howell claimed. While it is difficult to predict where a language is heading, the popular vocalic pronunciations of words such as *milk*, and *pill* suggest that its days are numbered after /ɪ/. L-vocalisation, as it is known, is sweeping the Anglophone world, including Australia and New Zealand.

The /l/ allophony is triggered by syllable structure. Other cases of complementary distribution may depend more on the adjacent sounds. We will consider two more examples, this time from Japanese and German, and we will look more closely at the way to go about determining phonemic status.

Consider the twenty words from Japanese listed in example 8.4 (this time unsorted), and see whether you can figure out whether the sounds [t], [ts] and [tʃ] represent one phoneme, two phonemes or even three phonemes.

(8.4)

1 [te] ‘hand’

2 [tsuri] ‘fishing’

3 [kutʃi] ‘mouth’

4 [geta] ‘wooden clogs’

5 [ta] ‘rice field’

6 [tʃi] ‘blood’

7 [ita] ‘board’

8 [tokoro] ‘place’

9 [to] ‘door’

10 [mitsu] ‘three (things)’

11 [tsugi] ‘next’

12 [tʃidʒin] ‘aquaintance’

13 [kutsu] ‘shoe’

14 [atsusa] ‘heat’

15 [takai] ‘high’

16 [hatʃi] ‘eight’

17 [tʃisei] ‘topography’

18 [tegami] ‘letter’

19 [utʃi] ‘house’

20 [tʃitʃi] ‘father’

To qualify as allophones of the same phoneme, the appearance of the sounds has to be predictable; therefore, we need to check the distribution of these three sounds and see whether there is something systematic about the environments in which they occur. If their environments don’t overlap (in other words, if they are in complementary distribution), this means their environments are predictable, and we know they are not distinctive but variants of the same phoneme. If there is an overlap in the environments, though, then the sounds are distinct.

In tackling a problem like this, there are two simple steps to follow:

Step 1 Check for minimal pairs.

Step 2 Check where these sounds occur (including sounds before and after).

A look through the data tells us that there are no minimal pairs, so we move onto Step 2. For this step, make a chart that lays out the environments of the sounds in question for all words (see Table 8.8). Here there are a couple of useful conventions: ‘#’ signifies a word boundary and ‘_’ signifies where the phone appears (for example, #_e means ‘word-initially and followed by e’).

Table 8.8 Distribution of [t], [ts] and [tʃ] in Japanese data from example 8.4

[tʃ]
#_i u_e (6, 2, 11, 19, 8)
#_u i_i e_a (1, 4, 10, 12, 17, 20)
a_a #_i (1, 5, 6, 11, 15)
i_u i_a (20, 7, 3)
a_o #_u (1, 4, 8, 9)

A chart like this sets out clearly the environments for each sound. And on this basis you can make the following generalisations:

- [ts] occurs before [u]
- [tʃ] occurs before [i]
- [t] occurs everywhere except before [u] and [i].

The sounds are in complementary distribution, so we know they are allophones of the same phoneme. We can move onto a new step.

Step 3 Select the phoneme.

As we said, it doesn't really matter what sound you select to represent the phoneme (and put in the slashes). You could put the image of a flower-pot in there if you like – the phoneme is an abstract entity (it's the allophones that are real). Nonetheless, it makes sense to select the phone that has the widest distribution – the 'elsewhere' phone – and clearly [t] stands out as this one. (It also makes sense because speakers would automatically choose the most frequent sound – the one they 'know' – especially if it's a language with a strong culture of literacy). So, we can model the relationship between phoneme and its allophones as we did earlier (see Figure 8.4).

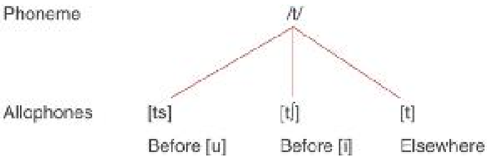


Figure 8.4 The Japanese phoneme /t/

There is a next step.

Step 4 Describe what is happening to these sounds.

This next stage can be done in prose or a more formal notation, as shown in example 8.5 (we recommend you try both). Here again, there are a couple of handy conventions: an arrow to indicate the process, and a long slash '/' (which means 'in the environment of'). As in Table 8.8, the dash '_' indicates the position of the allophones relative to the adjacent sounds.

(8.5)

/t/ → [ts] / _ [u]
[tʃ] / _ [i]

In prose, the phoneme /t/ becomes [ts] when it is in the environment of a following [u] and [tʃ] in the environment of a following [i]. You don't need to state the 'elsewhere' environment because it can be assumed (but sometimes it's useful to make it explicit – in other words, [t] everywhere else).

The triggering process here is assimilation; in both cases, the articulation of the alveolar consonant has shifted in anticipation of the (high) quality of the following vowels (compare what happens to the consonants in the English word: *create–creation–creature* and *electric–electrician–electricity*).

The shape of English loans in Japanese

Japanese is something of a 'vacuum cleaner of language' (recall that this was David Crystal's description of English given in Chapter 4). Like English, Japanese has filched vocabulary from many other languages; indeed, there are entire dictionaries of *gairaigo* 'loanwords'.

Borrowings can go two ways: they can either be adopted into a language without being integrated (keeping something like their original form); or they can be integrated into the phonological system of the borrowing language (adapted into a new shape to suit). What you typically find is that languages are predominantly adopters or predominantly adapters. Japanese and English are typical adapters – they alter loans to conform to their phonology. Here are some English loanwords that reflect the facts of Japanese phonology just discussed. (Note, both *ch* and *cc* = [tʃ].)

- *chiketto* 'ticket'

- *dasutobokkusu* (< dust box) ‘rubbish bin’
- *fecchi* ‘fetch’ (computer jargon)
- *hottokēki* ‘hot cake’ (a type of pancake)
- *pantsu* (< pants) ‘underpants’
- *rīchi* (< reach) ‘close to winning (a game)’
- *risutora* (< restru(cturing)) ‘downsizing’
- *toraburu* ‘(legal) trouble’
- *waishatsu* (< white shirt) ‘dress shirt’

Living up to its vacuum-cleaner status, English occasionally reborrows expressions: *anime* derives from the English *animation* but has come back into English with the specific meaning ‘Japanese animation’, and *kosupure* (< cos(tume) play) is used to describe anime outfits. Professional wrestling fans now use the Japanese term *pororesu* (< professional wrestling).

As a final example, consider the examples from German in example 8.6 and see whether you can determine the phonemic status of the velar fricative [x] and the corresponding palatal [ç]:

(8.6)

1 *Dach* [dax] ‘day’

2 *Dächer* [dɛçə] ‘roofs’

3 *Loch* [lox] ‘hole’

4 *Löcher* [løçə] ‘holes’

5 *hoch* [hox] ‘high’

6 *höchst* [høçst] ‘highest’

7 *Buch* [bux] ‘book’

8 *Bücher* [byçə] ‘books’

9 *lachen* [laxən] ‘to laugh’

10 *lecheln* [lɛçəln] ‘to smile’

11 *Lauch* [laux] ‘leek’

12 *nicht* [nɪçt] ‘not’

13 *Acht* [axt] ‘eight’

14 *sich* [zɪç] ‘self’

15 *Sprachen* [ʃpraxən] ‘languages’

16 *reich* [raɪç] ‘rich’

In example 8.6, [x] and [ç] are in complementary distribution. A table detailing the phonetic context (Table 8.9) reveals this straight away. This time, the trigger environment has to do with the preceding vowel (notice they never overlap).

Table 8.9 Distribution of [x] and [ç] in German data from example 8.6

[x]	**[ç]**
a_# (1)	ɛ_ə (2, 10)
o_# (3)	ø_ə (4, 5)
ɑ_s (6)	ø_# (7)
ʊ_ə (8)	y_ə (9)
aʊ_# (11)	ɪ_# (12)
a_t (13)	ɪ_# (14)
a_ə (15)	aɪ_# (16)

Again, the process is assimilation. In this case, the velar variant (the ‘backer’ version) [x] occurs after back vowels and the palatal (the ‘fronter’ version) [ç] is found after front vowels.

(8.7) /x/ → [ç] / [front vowels] _

In other words, /x/ becomes [ç] after front vowels. Remember, it doesn’t really matter which phone is chosen to represent the phoneme – you just change your rule accordingly; if we selected the other variant, we would simply alter the rule to:

(8.8) /ç/ → [x] / [back vowels] _

So-called ‘free variation’

One situation we haven't yet considered is where sounds appear in the same context but without causing a change in meaning, so they are not contrastive. In this case, they are allophones of the same phoneme, but are said to be in **free variation**. Consider the example we used at the start of this chapter – the different pronunciation of English /t/. As noted above, we can describe it as a phoneme with a number of allophones, as shown in Table 8.10.

Table 8.10 Allophones of English /t/

[illegible]
[illegible]
[illegible]
[illegible]

These are all phonetically similar sounds that occur in complementary distribution: aspirated [tʰ] is found word-initially; unaspirated [t] follows [s]; unreleased [t̚] is found word-finally; and flapped [ɾ] appears between vowels. This can be shown diagrammatically, as in Figure 8.5.

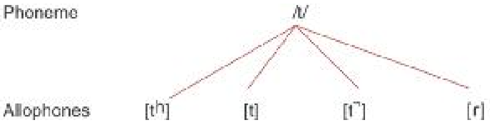


Figure 8.5 The English phoneme /t/

We can state the distribution of these variants more formally as shown in example 8.9:

(8.9) /t/ →

[tʰ] / #_ vowel
[t] / [s] _ vowel
[t̚] / _ #
[ɾ] / vowel _ vowel

The variation we have seen up to this point is conditioned. As in the previous examples discussed here, there is some kind of linguistic trigger involved; however, it is also possible for variation to occur that is not conditioned linguistically. It may be that individual speakers select differently depending on the occasion, or that the allophones selected may vary from speaker to speaker.

In pronouncing the /t/ word-finally in words like *bet*, a speaker may sometimes (in more careful articulation) actually release the stop, even to the extent that there might be aspiration. So, *bet* has potentially three different stops word-finally: unreleased [t̚], released [t] and aspirated [tʰ]. Even though the phones are found in the same context, they do not signal a difference in meaning. They can't represent different phonemes, but are allophones of the phoneme /t/ in free variation.

(8.10) Free variation of allophones: [bet̚] ~ [bet] ~ [betʰ]

Of course, 'free' is a misleading description of this variation, since it isn't really free at all – there is a whole heap of sociolinguistic significance attached to these variants. It might be variation associated with occasions of use (for example, how (in)formal the situation is), or it might be variation associated with speakers: including regional variation (the geographical background of speakers; social variation (the age, gender, sexual orientation; socioeconomic class and education of speakers); and even personal variation (the personality, physique and physical condition of individual speakers).

Other variants of /t/ have clear regional significance. Think of the 'glottal' stop pronunciation of words like *butter* by Cockney English speakers [bɐʔə]; there is also a glottalised variant in Australian and New Zealand English, but it more usually occurs in medial contexts (*cutlass*) and end positions (*shut*), and less so between vowels (as in *butter* and *get out*). Some Australian English speakers also produce something called a fricated /t/ (the symbol is [tˢ], which we touched upon in our discussion of forensic linguistics in Chapter 7). This is especially frequent in intervocalic and word-final positions (for example, *That's a beautiful* cat [kætˢ]), and is associated with women, the middle class and cultivated vowel sounds (some have attributed this feature to Irish English). So, the term 'free variation' is confusing – there are constraining variables (as we explore in Chapter 10), but they don't happen to be phonological.

Phonetic similarity

In English, [h] and [ŋ] are actually in complementary distribution: [h] only ever appears at the start of syllables (as *in hat, ahead, apprehend* and *vehicular*, for example), whereas [ŋ] can only ever appear at the end of a syllable, either by itself (as in *ring, singer, long*), or before a velar stop (as in *finger, longer, sink*). There are no minimal pairs, or even near-minimal pairs, to distinguish these. But, do we want to call these allophones of the same phoneme?

All the examples of allophones we have seen so far have involved sounds that are phonetically similar, to the extent that native speakers usually perceive them to be the same sound (aspirated and unaspirated /t/, or 'clear' and 'dark' /l/, for instance). It is hard to imagine, however, that speakers would ever link [h] and [ŋ] in this way. So, we do not want to say that when sounds occur in complementary distribution it is automatically a sign of allophony; rather, we put a condition in place that says for sounds to be considered allophones of the same phoneme they have to be phonetically similar. In the case of [h] and [ŋ], they are both consonants, and that is all they share! Consider them in terms of the three parameters of description we gave earlier (voicing, place and manner of articulation):

(8.11)

[h]	[ŋ]
voiceless	voiced
glottal	velar
fricative	(nasal) (stop)

It is hard to fathom that a single phoneme would have such different pronunciations. The principle of phonetic similarity thus allows us to say that /h/ and /ŋ/ are clearly separate phonemes. In Chapter 11, we will see the historical reasons for why these two phonemes are so curiously restricted (sometimes dubbed 'defective distribution'). It happens that /h/ is an endangered phoneme, and probably would have disappeared long ago if writing hadn't thrown it a lifejacket. On the other hand, /ŋ/is a new phoneme (historically, it was an allophone of /n/); it hasn't yet had the chance to spread its wings! Another example of an English sound with a defective distribution is /ʒ/; it doesn't occur very often in word-initial position (which is why English speakers have such trouble with French borrowings such as *jus* and *genre*). This is another new phoneme, which grew out of the sequence [zj] in words such as *measure* and *pleasure*, and is supported by French borrowings such as *beige* and *rouge*.

In sum, we can say that phones have to be phonetically similar to each other to be considered allophones of the same phoneme. Highly dissimilar sounds are separate phonemes, even if their distribution is complementary (like /h/ and /ŋ/). Phonetic similarity is based on the idea of shared features, but it is still rather vague – how similar is similar? We can fine-tune this notion by looking more closely at this idea of shared features, which we will do in a moment. But first, we want to look at what are hopefully some helpful illustrations to cement the notions of phoneme and allophone.

Metaphors to help illuminate the phoneme

Beginning students of linguistics are sometimes confused by the concept of the phoneme, particularly notions of complementary and contrastive distribution, and over the years all sorts of real-life analogies have been suggested to help make these rather abstract concepts a little more tangible. Back in the fourth century bce, the Greek philosopher Aristotle made much of the instructive value of metaphors through vividness and novelty of expression. Metaphors highlight reality and also generate aspects of reality that go beyond literal language. They offer two loaves where there seems to be one – they might even throw in a fish (and to understand this image, you need to visit the quotation we used to frame our discussion of metaphor in Chapter 6). We hope that at least one of the following analogies – whether it involves novels, superheroes or cane toads – will give you an insight into how phonemes and allophones work.

One way to think about this is to compare phonemes to classes of books and allophones to the individual books. If you visit a library, for example, you will find whodunnits, Westerns, science fiction, romance novels, and so on. Perhaps you select a whodunnit (compare the phoneme /t/), which is of course not a single book but a class to which a number of single books belong, including *The Hound of the Baskervilles, A Study in Scarlet* and *The Daughter of Time* (compare allophones, such as [tʰ], [t] and [t̚]). This image rather nicely captures the idea of the phoneme as a class or category to which certain sounds (or phones) belong; so, a phoneme is then an abstract unit in the sound system (like the category of whodunnit in the library).

A popular analogy for complementary distribution (though we're not sure who thought of it first) is 'Clark Kent' and 'Superman'. These two characters never appear in the same environment: Superman appears in the environment of a crisis; Clark Kent is never in the environment of a crisis. Thus, these two characters are in complementary distribution, and therefore are different realisations of the same person, just as allophones are different forms (or realisations) of the same phoneme.

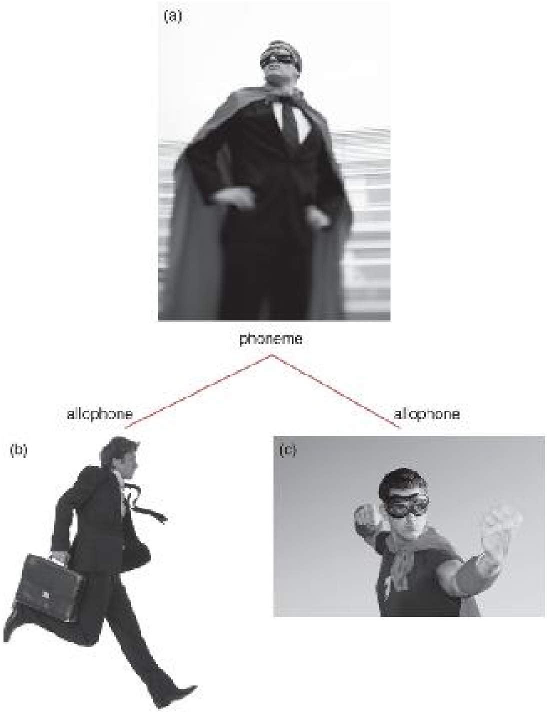


Figure 8.6 ‘Clark Kent’ and ‘Superman’ are in complementary distribution

Terry Crowley and his colleagues offer another novel metaphor that explains notions of complementary and contrastive distribution in terms of the cane toads and the cane beetles of Far North Queensland. Cane toads were introduced to Australia to wipe out the cane beetle, whose larvae were eating the roots of sugar cane and killing or stunting the plants. In 1935, more than 3000 cane toads were released into the sugar cane plantations. The plan was a failure because, like aspirated and non-aspirated consonants in English, cane toads and cane beetles are in complementary distribution – cane toads stay near the ground while the beetles live in the upper stalks of cane plants.

Finally, if novels, superheroes or cane toads don’t do it for you, our colleague Simon Musgrave has his own rather ingenious analogy of two identical twins in the same play (compare two sounds). It is only when the two players appear on stage at the same time that they contrast and you realise that they are distinctive individuals (in other words, separate phonemes).

8.4 Distinctive features and natural classes

In Chapter 6, you saw how the meanings of words could be broken down into simpler meanings by a fixed set of binary ([+] or [–]) features. Recall that words sharing one or more components are said to belong to the same semantic field. For example, *woman* and *girl* are two words from the semantic field of ‘the human race’: *woman* [+ HUMAN, + ADULT, + FEMALE] and *girl* [+ HUMAN, – ADULT, + FEMALE]. These two lexemes contrast only in the feature [± ADULT].

The primary focus of phonology is to uncover how sounds function and how they are patterned in a given language. It is often useful to break down sounds in this same way into their basic characteristics to understand how phonological systems in different languages work. To some extent we have already been doing this when we described sounds in terms of the three parameters of description: place of articulation, manner of articulation and voicing. For example, the consonants /s/ and /z/ contrast with respect to only one feature, [± voice] (see example 8.12).

(8.12)

/s/	/z/
+ fricative	+ fricative
+ alveolar	+ alveolar
– voice	+ voice

Phonologists have worked on refining these features to come up with the most economical set and a number of different distinctive feature systems now exist. Here we will just give you taste of some of the features that have been proposed – enough that you will be able to group sounds together and to see what links sounds that behave similarly with respect to the phonological processes described earlier. Some of these features will already be familiar to you.

As suggested by the descriptions in Table 8.11, distinctive features can be used to set up **natural classes** of phonemes and this is helpful when it comes to expressing phonological processes.

Table 8.11 Some distinctive features

[± voice]	Vibration of vocal cords (distinguishes voiced from voiceless consonants)
[± consonantal]	There is contact or near contact at some point of articulation (distinguishes stops, fricatives, nasal, liquids from vowels)

[± nasal]	Soft palate is lowered, resulting in nasal resonance (distinguishes oral stops from nasal stops)
[± lateral]	Air proceeds along the sides of the tongue (distinguishes lateral liquids from rhotic liquids)
[± sonorant]	Characterised by full resonance (unites vowels, nasals, liquids and glides, and distinguishes them from stops, affricates and fricatives)
[± coronal]	The blade or apex of the tongue is raised (distinguishes dental, alveolar, (alveo)palatal and retroflexed sounds from bilabial, velar, uvular and glottal sounds)
[± continuant]	The oral cavity is not blocked off completely (includes vowel, glides, liquids and fricatives; used to distinguish stops from fricatives)
[± strident]	There is high-pitched turbulence characteristic of some fricatives and affricates (distinguishes 'hissy' fricatives, or sibilants, such as /s/ from others, such as /θ/)
[± high]	The body of the tongue is raised (high vowels, glides, (alveo)palatal and velar consonants)
[± back]	The body of the tongue is retracted (moved back) (back vowels, velars, pharyngeal consonants)
[± round]	Articulated with lip-rounding (unites labial and labialised consonants and rounded vowels)

We might have a process that applies to all the nasal consonants and nasal vowels in a language; rather than write separate rules for each of the sounds affected, we can unite all the relevant segments with the feature [+ nasal]. Or consider the palatalisation of the German velar fricative described above. It involves the consonant taking on the tongue position of the preceding vowel and we can capture this with features. Our earlier rule could be rewritten this way:

(8.13) /x/ → [– back] / vowel [– back] _

Recall how English laterals disappeared between back vowels and bilabial and velar consonants in words such as *walk, talk, balm* and *calm*. These consonants do not involve any movement of the flexible part of the tongue and are non-coronal. In other words, we can unite the consonants that helped trigger the demise of /l/ with the feature [– coronal].

Natural classes of phonemes can often be characterised by a small number of features. For example, [b, d, g] form a natural class of voiced stops in many languages. They can collectively be characterised as [– continuant, + voiced], and this set of features is enough to distinguish them from other sets of sounds. On the other hand, [f] and [n] do not have much in common, so do not readily form a natural class. Few, if any, phonological rules apply to them both and, appropriately, it is impossible to refer to these sounds and no others in a single feature matrix.

8.5 The pronunciation of morphemes: Morphs and allomorphs

Recall in Chapter 5 that morphemes were described as often having more than one shape and these various shapes are called **allomorphs**. Things often happen to morphemes when they combine to form words; often these allomorphs are simply alternative pronunciations (phonetic forms) of morphemes. In Chapter 5, we provided the example of forming the plural in English. Speakers of English select between alternative forms, depending on the voice quality of the preceding sound. This is shown in Table 8.12.

Table 8.12 Consonant assimilation in English (plural formation)

Plural	**[s]**	**[z]**	**[əz]**
	[læfs] ‘laughs’	**[bægz] ‘bags’**	[dʒʌdʒəz] ‘judges’
	[kʌps] ‘cups’	[klʌbz] ‘clubs’	[spaɪsəz] ‘spices’
	[bits] ‘bits’	[bɜdz] ‘birds’	[tʃizəz] ‘cheeses’

The plural for each of these words is written as <(e)s>, but this has three very different pronunciations: the first is a voiceless sound, [s]; the second is a voiced sound, [z]; and the last has a vowel before it, as in [əz]. So, the plural suffix in English has three alternative realisations, or allomorphs. The final segment of the stem determines the form of the allomorph. As we noted in Chapter 7, this rule guides native and fluent speakers of English to subconsciously and automatically select the correct plural for new nouns such as *grotch, wug* and *sputnik*.

(8.14)

Allomorph	**Environment**
[s]	Occurs after stems ending in a voiceless segment (excluding sibilants).
[z]	Occurs after stems ending in a voiced segment (excluding sibilants).
[əz]	Occurs after stems ending in a voiced/voiceless sibilant.

We describe these alternations just as we did in the case of allophony, by outlining the contexts within which variants occur (that is, their distribution). The only difference is that we are analysing phonetic variants of morphemes, not phonemes. Just as we did for allophony, we can describe the distribution either in words (as we have just done), or by using a more formal notation. When we do this, we select one form as basic and write the rules to derive all other forms; the form to choose is the one that makes your rules the most straightforward. Let's say we choose the suffix /z/.

Plural formation in English

/z/ → [əz] / [sibilants] _

[s] / [voiceless non-sibilant consonants] _

Similar conditions hold for the past tense of most verbs (see Table 8.13).

Table 8.13 Consonant assimilation in English (past tense formation)

Past	[t]	[d]	[əd]
[illegible]	[tæɡd] 'tagged'	[illegible]	
[tæpt] 'tapped'	[illegible]	[illegible]	
[reɪst] 'raced'	[reɪzd] 'razed'		

These allomorphs are phonologically conditioned because we can describe their distribution – the environments in which they occur – in phonological terms. The distribution statement for the past tense suffix (see example 8.15) is very similar to that of plural formation.

(8.15)

Allomorph	Environment
[t]	Occurs after stems ending in a voiceless segment (excluding alveolar stops).
[d]	Occurs after stems ending in a voiced segment (excluding alveolar stops).
[əd]	Occurs after stems ending in an alveolar stop.

These are examples of phonologically conditional allomorphy. They occur because of phonological processes such as assimilation (in this case of voice) and epenthesis (vowel insertion); both processes make pronunciation easier. The allomorphs are in complementary distribution (that is, each occurs in certain environments in which the others never occur).

There are many different ways in which you can write these phonological rules. You might draw on the distinctive features we discussed earlier to better reveal the trigger for the alternative pronunciations. At this stage, it doesn’t really matter, as long as you capture the allomorphs and state where they occur. Try your hand at writing a rule for past tense formation; it would be similar to the rules that can be used to describe the distribution of the plural.

8.6 Prosodic phonology

In Chapter 7, you were introduced to the **prosodic** (also called **suprasegmental**) features of sounds, such as stress and tone. These are the properties that stretch beyond segments (vowels and consonants) and involve syllables or words. They too can be part of the phonology of a language.

Syllable structure

> Everyone knows that ‘a syllable is what syllable has three of’.
>
> Roger Lass (1998), *Phonology: An Introduction to Basic Concepts*

We established in the last chapter that there is plenty of evidence for syllables being psychologically real for speakers. Recall also that syllables comprise three or four different parts. The onset and the rhyme are the main two parts, but the rhyme can also be divided into a nucleus and (optionally) a coda. The words *snap* and *snide* share the onset /sn/, but have different rhymes (/æp/ and /ɑed/), while the word *at* only has a rhyme /æt/. The nucleus is the vowel (or diphthong) in these examples /æ/ or /ɑe/, and the coda is the final consonant or consonant cluster, in these examples, /p/, /d/ or /t/. The word *sniffle* has two syllables [snɪf] and [l̩]. The second syllable does not have an onset and the rhyme only contains the nucleus [l̩].

Suppecruppet Luppanguppuage

All over the world, secret languages exist that draw on syllable structure to distort words in some way (typically, they serve a dual purpose of secrecy and solidarity, preventing bystanders from understanding what is being said and serving as a 'clique' device). Many shove infixes inside syllables. 'Upp-Upp' language, for example, inserts *upp* before the rhyme of each syllable of a word (that is, the vowel(s) and any consonant(s) that might follow). 'Hello, how are you' becomes *huppelluppo, huppow uppare yuppou*. To which the reply might be: *Vupperuppy guppood thuppank yuppou*. There are many varieties out there that draw on different infixes; some of you will be familiar with *iz* or *izn* from hip-hop slang (for example, *hizouse* 'house' and *shiznit* 'shit'). Or try adding 'lf ' after the nucleus of the first syllable, then repeat the vowel(s) to finish the word. 'A double barrelled shotgun' becomes a *dulfubble balfarrel sholfotgulfun*.

Some secret languages can be mind-bogglingly complex. One Portuguese variety requires no less than four different manoeuvres: add [p] at the end of each syllable; after the [p] you've just added, copy the rhyme of that syllable; change all high vowels to low vowels but only in open syllables (so syllables ending in a vowel); and finally, disregard the original stress and instead stress the copy of each rhyme. Of course, do all this at normal conversational speed!

These secret languages provide excellent evidence for the psychological reality of the syllable. Small children can speak Upp-Upp and similar languages fluently (and often do, much to the irritation of parents who can't follow what is being said). Yet, these children wouldn't be able to explain what a syllable is, let alone where the onset, nucleus and coda of syllables occur. It is an impressive performance, but we don't usually think of it this way because – as for so much of our language – we just take it for granted.

Phonotactic constraints

> Don't know what a Schelgesetzentwurf is, but I keep as excited over it and as worried about it as if it were my own child. I simply live on the Sch.; it is my daily bread. I wouldn't have the question settled for anything in the world.
>
> Albert Bigelow Paine quoting his subject, in *Mark Twain: A Biography, 1835–1910, complete*

Rules that describe the possible syllable structures are called

phonotactic constraints. In English, there are restrictions – for example, on the combinations of consonants that make up a two- and three-consonant cluster. The following list shows the possibilities for three-consonant clusters in a syllable onset at the beginning of a word.

/s + p + l, ɹ, j/

/s + t + ɹ, j/

/s + k + l, ɹ, j, w/

Try to think of words to illustrate each of these possible consonant combinations (you will find that some of the combinations, such as /skl/, are quite restricted.)

What’s in a name? Curious sound patterns

Personal names can often have odd phonological structures. The cluster [gw], for instance, is extremely rare in English – it appears in only a handful of mainly obsolete curiosities, including *gwan* (a condensed form of ‘go on’) and *gwash* (a version of the French *gouache* ‘a method of painting’). But, it is alive and well in *Gwendolyne*. While [dw] is not quite as rare, most words are either obsolete (*dwine* ‘to pine away’), or are derivations from *dwarf* and *dwell* (which makes slang *dweeb* an interesting recent addition). It is not surprising that the personal names *Dwight* and *Dwayne* are often pronounced with an epenthetic vowel inserted between [d] and [w], to make the cluster a bit more agreeable [dəwæɪn]. Another source of weird phonology, as we saw in Chapter 7, is interjections, the noisy responses speakers use to express some sort of emotion (like *ugh* and *humpf*). Often these have unusual phonotactics. Something like *psst*! (used to signal attention) is unusual in that it has a syllabic fricative (the fricative takes up the nucleus of the syllable).

Because languages can have such different phonotactics, loanwords are often adapted (sometimes spectacularly) to conform to the structural constraints of the borrowing language phonology. They might use deletion strategies to break up offending consonant clusters, or epenthesis strategies (vowel insertion) to adapt the clusters. If you revisit the earlier list of English loanwords into Japanese, you will see that epenthesis is the usual way of adapting these loans to Japanese phonotactics (for example, *Christmas* → [kuɾisumasu]). English loanwords into Māori are similarly altered. English has phonemes without Māori equivalents, and Māori has no consonant clusters and no closed syllables (every consonant must be following by a vowel). As Table 8.14 shows, Māori uses a combination of deletion and epenthesis: vowels are added to loans ending in a consonant (often, although not necessarily, the same as the second-last vowel of the loanword); initial /s/ clusters drop the sibilant; other epenthetic vowels are added to break up clusters (and there are interesting conditions determining what these are).

Table 8.14 English loanwords into Māori

English source	**Māori loanword**
blanket	*paraikete*
cricket	*kirikiti*
taxi	*takihī*
hospital	*hōhipera*
strawberry	*tōpere, rōpere*
street	*tiriti*
slipper	*hiripa*
blackberry	*parakipere*
block	*poraka, paraka*
blue gum	*pūrukamu*
scholarship	*karahipi*
lunch	*rānihi*

To sum up, phonotactic constraints describe the permissible sequences of vowels and consonants in a syllable. Thus, they identify the thousands of words that do not exist but could from the thousands more that cannot because they would not have permissible syllable structures.

Prosodic features

Let’s now consider how prosodic features relate to other parts of phonological systems. **Length** (or **quantity**) relates to the duration of sounds. You have already seen that vowels can be short and long (indicated by two dots). This fact is of limited interest to phonology, unless languages use this feature to distinguish words. German is one such language. Consider the minimal pairs in Table 8.15.

Table 8.15 Long versus short vowels in German

Rose ‘rose’	*Rosse* ‘steeds’
beten ‘to pray’	*Betten* ‘beds’
Ruhm ‘glory’	*Rum* ‘rum’

These minimal pairs confirm that vowel length is phonemic in German. Vowel length is unpredictable and, although it interacts with some aspects of vowel quality (such as the tense/lax distinction), it is largely length that distinguishes these words (long vowels are about twice as long as short vowels).

Some languages also use long and short consonants to distinguish words. In Japanese, consonant length is distinctive (for example, *bagu* '(computer) bug' versus *baggu* 'bag'). In Old English, consonant length was also phonemic, as shown by minimal pairs such as *raca* 'rake' versus *racca* 'cord (part of a ship's rigging)' (the /kː/ here is pronounced much like the [k] in the middle of a compound word such as *book-keeper*). Modern English no longer has distinctive long consonants (the double consonants you find in words like *matter, dropped* and *tunnel* indicate that the preceding vowel is short).

Recall from Chapter 7 that **stress** is the relative emphasis that can be given to certain syllables. In some languages, this is predictable. Māori has a regular rule whereby the main stress falls on the initial syllable of a word; exceptions occur where there are long vowels (these attract stress) and where the stress falls more than four syllables from the end of the word. So, in this language, stress is not phonemic – we can never find a minimal pair that involves the position of stress.

Other languages, such as English and Russian, have variable stress; in other words, the location of the stressed syllable can't be predicted by a general rule of stress placement (as in Māori), but must be learned as part of the pronunciation of words. In these languages, stress can distinguish otherwise identical words and is therefore phonemic; for example, INvalid 'patient' versus *in*VAL*id* 'illegitimate'. As touched on earlier, stress is also used to signal the difference between English noun and verb pairs – nouns are stressed on the first syllable and verbs on the second:

They always make good PRO*duce.* (noun – stress on first syllable)

*They always pro*DUCE *excellent wine.* (verb – stress on second syllable)

To conclude this section on suprasegmentals, we revisit **tone**, a feature that relates to the pitch contours of a syllable. When languages use this feature to distinguish words from each other, we describe them as **tonal languages**. Turung is one such language; it has three tones.

In tonal languages like Turung, pitch is phonologically significant – as Table 8.16 shows, sequences of segments have a totally different meaning depending on the pitch. Contrast this with a language like English, where pitch (together with other features) serves only to signal stress.

Table 8.16 Tone in Turung

1	Falling	*wa*^{21}	'definite'
2	High rising then falling	*wa*^{22}	'bite'
3	Level	*wa*^{23}	'make (a basket)'

8.7 Representing spoken language

We started this chapter by pointing out the importance of phonology when it comes to devising orthographies (ways of capturing spoken language). We finish the chapter by looking briefly at different writing systems and how they evolved.

Before writing, experiences were recorded in pictures (pictographs). Like the international prohibition sign ⃠, these are pictures conveying messages. However, they don't stand for linguistic units (words or morphemes), so aren't yet actual writing – for this we need spoken words to be analysed into segments. Here are the three main ways to do this (although writing systems can combine them):

1 *logographic system* – a symbol for each word (or morpheme); for example, the 'characters' of Chinese and the Japanese *kanji* (English examples of logographs include 3 for 'three' and $ for 'dollar')

2 *syllabic system* – a symbol for each syllable; for example, Cherokee and Japanese *katakana* for foreign language words and *hiragana* script for native words; (English examples include *ok* standing for the syllables [o] and [ke] and *bbq* for *barbecue*)

3 *alphabetic system* – each symbol represents the separate sounds that make up a syllable; for example, English and Russian.

Manipulating scripts in Japanese

Japanese writing is famous for its combination of four distinct scripts: the *kanji, hiragana, katakana* (as just described), and *rōmaji* (basically the Roman alphabet used for advertising, signage and brand names). Wes Robertson is someone who has investigated how Japanese writers can convey different types of emotional and semantic information by drawing on these different scripts. In a study of script selection in manga comics, he shows writers manipulating how pronouns are represented to index different identities for their characters; for example, (male-exclusive) first person pronoun *boku* in katakana script (ボク) conveys a 'semi-polite but childish or show-offy' feel versus the 'semi-polite but more adult feel' of the kanji script (僕). This flexibility of orthography is an important way of creating meaning in written Japanese: you might compare how English writers select words, spelling variants and different styles of writing to capture different nuances of meaning (recall how umlauts are used in the names of heavy metal bands like Motörhead).

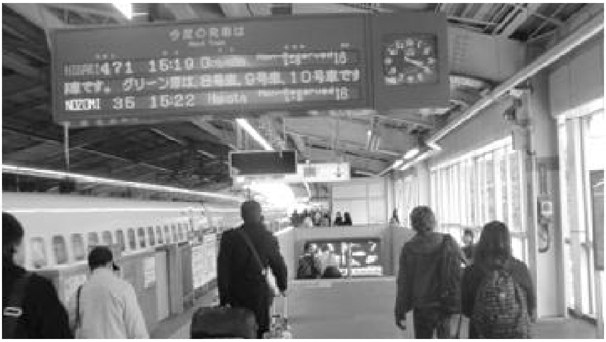


Figure 8.7 Sign at a railway station showing the use of mixed scripts in Japanese writing

The development of writing

The earliest writing was logographic. Typically, logographs start life as pictographs, but with time they become more abstract. The following example shows how the logograph for 'head' evolved into the wedge-shaped symbol used in the Sumarian script (known as cuneiform):

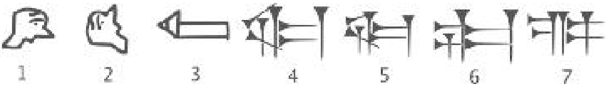


Figure 8.8 The evolution of the Sumerian logograph for *sag* 'head'

It isn't easy to draw pictures, particularly of abstract concepts, but people figured out that they could draw a picture of words that sound the same. So, you find that some logographs extended to homophones of the original word symbolised; compare the English 'for sale' sign that uses *4* for *for*. This is known as the **rebus principle** (a kind of punning). These days it is associated with texting (e.g. *8* for 'ate' as in *activ8*). Here is a famous rebus children's rhyme:

2 Y's U R

2 Y's U B

I C U R

2 Y's 4 Me.

Once logographic symbols are associated with phonetic form, we see the beginnings of a syllabary (one early syllabary is 'Linear B', the script of Mycenaean Greek from around 2000 bce).

The earliest known system approximating an alphabet was North Semitic (developed around 1700 bce in Palestine and Syria). This had 22 consonant symbols (vowels were largely predictable). This system gave rise to the Hebrew, Arabic and Phoenician writing systems. The Greeks then took the Semitic scheme and systematically added symbols for vowels. Note that the Semitic writing goes from right to left and the Greek changed it to a left–right sequence (an intermediate stage involving both directions was called *boustrophedon* 'in the manner of the ox turning' – imagine an ox ploughing a field). This writing system then spread to Italy (around 800 bce) and was adopted by the Etruscans, who then passed it to the Romans. Out of this grew all the Western alphabets (including the runic and Cyrillic systems).

The Germanic tribes who first came to Britain in the fifth century didn't write much, but, when they did, they used the alphabet we now call the *futhorc* (named after the first six letters) or runic alphabet (from Old English *ru:n* 'mystery, secret'). The original 24 symbols were influenced by the Latin and Greek alphabet, but there were also some unique symbols (their angular nature suggests they were carved into wood or stone). You can see from the chart in Figure 8.9 that each runic symbol had its own name (unlike the Greek or Latin alphabets), and all but two names begin with the sound represented by the character (*eolh* and *Ing* represent [x] and [ŋ], and these sounds never occurred initially).

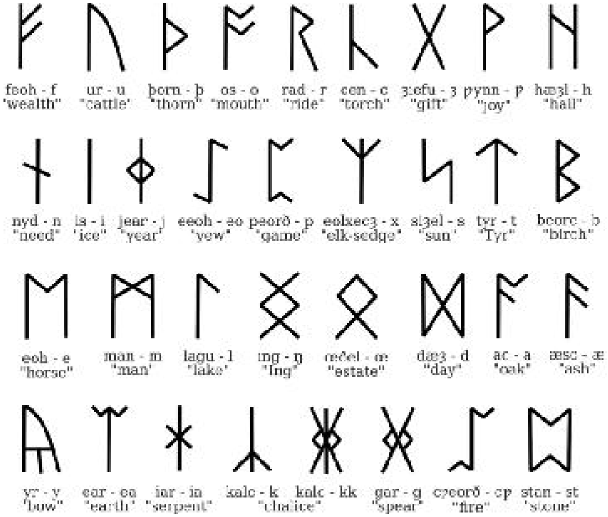


Figure 8.9 The runic alphabet of Old English

When the English converted to Christianity, they adopted the Latin alphabet (although clerical scribes continued to use some runes in their manuscripts). Eventually, with some twists and turns along the way, the letters of the alphabet assumed their current state and number.

Put your orthography where your head is

At the start of this chapter, we introduced the notion of a phonetic alphabetic system for English, and we showed what a bad idea it would be for an orthography to have all this phonetic detail. There is another drawback to respelling words more phonetically: it would mean we would lose one of the real virtues of English spelling. Take groups of words such as *electric, electricity* and *electrician*. Here we retain the root of the word in the spelling <electric>, but if these words were spelt on phonetic grounds, we would write <k> for the final sound in *elektrik*, <c> or <s> for *elektrisity* and <sh> for *elektrishan*. In other words, the meaning relationship between these words would be lost. English is full of examples like these. In the spellings of *create, creature* and *creation*, the shared root of these words is preserved, despite the addition of endings that change its pronunciation. Would people really want to spell *creature* as *creecha*? True, this might make life easier for learners of English as a foreign language, but for native speakers this is all predictable stuff. We don’t need to spell *creation* as *creashon*.

Current English spelling isn’t simply a defective phonetic script, but a more abstract system that has a combination of phonetic and lexical representation. This is a better reflection of what native speakers carry around in their heads. It is more psychologically real. As the American linguist Wayne O’Neil once said in defence of English spelling, ‘Don’t put your orthography where your mouth is; put it where your head is.’

Further reading

There are a number of excellent introductions to segmental and suprasegmental aspects of phonology; we recommend you start with Carr (1993), Lass (2000), Bybee (2003b), and Clark, Yallop and Fletcher (2007). Crowley et al. (1995) provide a memorable account of cane toads and cane beetles. Craig Evans (2013) has compiled a top ten list of famous trademark cases, and for interesting views on trademark issues generally, as well as a number of other areas of forensic linguistics, see Butters (2007, 2010). In a Language Log post, Victor Mair (2013) discusses English loans into Japanese, and for insights into English loanwords into Māori, we recommend the work of John C. Moorfield and Tania M. Ka’ai (2009). On aspects of phonetic interference, see Wells (2000). For accounts about writing see Allan (2010), Coulmas (2002) and Rogers (2004); see Robertson (2017) on scripts in Japanese. For early ideas on spelling reform, see Zachrisson (1931).

Exercises and discussion points

1 Comparison of English and Kriol

(a) In the following, compare the English words with the Kriol equivalents (Kriol is an English-based creole spoken in northern parts of Australia). What are the phonological processes that have led to the changes here? Where possible, give the linguistic label for this process (for example, assimilation, deletion or metathesis). If you cannot find an appropriate label, then simply describe the process.

Kriol	English	Translation

[illegible] ɔt]
[illegible] ɐːsk]
[illegible] iːp]
[illegible] ʉːn]
[illegible] ən]
[illegible] ðə]
[illegible] təkɐː]
[illegible] [bæg]
[illegible] ən]
[illegible] p]

Notes:* ‘j’ in Kriol represents a voiced palatal stop, [ɟ]; ** Kriol ‘r’ here represents a voiced flap.

(b) The English pronunciation here is typical of fairly careful speech. Identify six changes that occur in Kriol that also occur in conversational English. Give colloquial English examples to support your answer.

(c) Why do you think some Kriol speakers might insert *h* in the word *apinum* ‘afternoon’, pronouncing it as /hɐpɪnʉːn/? (Similarly ‘uncle’ becomes /hɐnkəl/ and ‘island’ is sometimes pronounced /hɑelənd/.)

2 Phonological processes – languages other than English

Name the phonological process involved in each of the sets of data below. Using both prose and formal notation, try writing the rules to express the processes.

(a) In Czech, certain consonants change before the diminutive suffix *-ek: pták* [pta:k] ‘bird’ versus *ptáček* [pta:tʃek] ‘little bird’.

(b) German: [buntə] ‘colourful’ (before the noun) versus [bunt] ‘colourful’ (after the verb); [bundə] ‘league’ (dative) versus [bunt] ‘league’ (nominative).

(c) Grammatical suffixes in Hungarian come in different forms, such as *-hoz/-hez/-höz* (roughly equivalent to the English preposition ‘to’); *a házhoz* ‘to the house’; *a gyerekhez* ‘to the child’; and *önhöz* (singular, polite form) ‘to you’.

(d) In Turkish, the stem changes when the accusative suffix is added: [dip] ‘bottom’ versus [dibi] ‘bottom (accusative)’; [sɛrit] ‘tape’ versus [sɛridi] ‘tape (accusative)’; and [grup] ‘group’ versus [grubu] ‘group (accusative)’.

(e) German: *Zaubererin* → *Zauberin* ‘sorceress’; *Superintendent* → *Superintend* ‘Superintendent’; *Bewundererin* → *Bewunderin* ‘admirer (feminine)’.

3 Turkish vowel harmony

In Turkish, vowels in suffixes take on certain features of the vowels that precede them. In the following data, identify and describe the suffixes, and determine the features to which their vowels are assimilating. (*Hint:* think about features such as backness, frontness, height and roundness of the vowels; it’s a good idea to plot all the vowels on a chart – it will help you to figure out how these vowels are harmonising.) Note that we have used the Turkish letters [ü] and [ö] to represent the rounded front vowels [y] and [ø].

Basic form		Suffixed form	
el	‘hand’	*elim*	‘my hand’
diş	‘tooth’	*dişim*	‘my tooth’
ev	‘house’	*evim*	‘my house’
gönül	‘heart’	*gönlüm*	‘my heart’
göz	‘eye’	*gözüm*	‘my eye’
baş	‘head’	*başım*	‘my head’
kız	‘girl, daughter’	*kızım*	‘my girl, daughter’
gül	‘rose’	*gülüm*	‘my rose’
kol	‘arm’	*kolum*	‘my arm’
tost	‘toast’	*tostum*	‘my toast’
eller	‘hands’	*ellerim*	‘my hands’
dişler	‘teeth’	*dişlerim*	‘my teeth’
evler	‘houses’	*evlerim*	‘my houses’
gözler	‘eyes’	*gözlerim*	‘my eyes’
güller	‘roses’	*güllerim*	‘my roses’
kollar	‘arms’	*kollarım*	‘my arms’
tostlar	‘toasts (slices of)’	*tostlarım*	‘my toasts (slices of)’
kızlar	‘girls, daughters’	*kızlarım*	‘my girls, daughters’

4 Transcribing oronyms

Because of the seamlessness of speech, sequences of sounds can be divided into words in more than one way; these are oronyms. For each of the following pairs, write out your pronunciation using the phonemic notation given in Chapter 7. According to your pronunciation of the strings, is each of these an oronym?

- **(a)** four candles – fork handles
- **(b)** grade A – grey day
- **(c)** It’s hard to recognise speech. – It’s hard to wreck a nice beach.

(d) The good can decay many ways. – The good candy came anyways.

(e) They have slain the Earl of Murray and laid him on the green. – They have slain the Earl of Murray and Lady Mondegreen.

5 Investigation of English stress

(a) In Old English, the root syllable took the major stress, and this is still the case in some Modern English words – for example, *FRIENDliness, LIKElihood, beCOMing*. The influx of loans has changed this considerably and there is now quite a lot of variation in English stress patterns. Give examples of as many different patterns as you can; if you include borrowings, check a dictionary to identify the source language.

(b) The words below show an earlier stress pattern (the vowel nucleus is given in upper case). Rewrite each word, indicating the stress pattern *you* use. In terms of syllable structure, what generalisations can you make about changes in stress patterns that have occurred?

chaRACter, HOSpitable, PEjorative, preCEdence, sonOrous, Applicable, conTEMplate, deCAdent

(c) Now that you've established the new pattern, where would you expect the stress to settle in the following controversial cases: *kilometre, controversy, comparable, metallurgy*.

(d) Some prescriptive texts allow only *KIlometre* and *CONtroversy*, and many argue that this is the only correct pronunciation. What comment would you make about this view?

(e) The following are also words currently in flux. Show where you put the stress in your speech. Where would you expect the stress to settle?

Caribbean, clandestine, irrevocable, lamentable

Research project

Exploring issues in phonology

Tips and resources

Research one of the following topics, then write an essay (approximately 1000 words). Your bibliography should have at least FIVE references.

1 Is the syllable as fundamental a unit in sign languages as it is in spoken languages? As part of your discussion, consider the evidence for this claim (for example, the ‘manual babbling’ stage in pre-linguistic development of Deaf children).

2 Dr Samuel Johnson once wrote, ‘Tongues, like governments, have a natural tendency to degeneration.’ Base your essay on this quotation. As part of your discussion, describe the phonological processes that might lead one to this conclusion, and (from a modern linguistic perspective) comment on the validity of Johnson’s observation.

3 English spelling is notoriously chaotic. Highlight some of the problems and discuss the kinds of changes that might be considered to make English spelling more rational. Give some indication of the linguistic problems associated with such changes. You may choose to include discussion of a specific attempt at spelling reform.

4 In Chapter 7, we described a distinction between stress-timed and syllable-timed languages to describe the rhythm that characterises these languages – French and Yoruba are said to be syllable-timed, whereas English and Russian are stress-timed languages. Give an account of the differences between stress-timed and syllable-timed languages (make reference to differences in syllable structure, vowel reduction and the phonetic realisation of stress).

For more activities, please go to www.cambridge.edu.au/academic/linguistics.

9

Syntax: The structure of sentences

◆

> I'm not sure what syntax means. But it must be bad because it's got 'sin' and 'tax' in it.
>
> William Rogers (American humorist)

9.0 Introduction

One of the key characteristics of human language that we identified in Chapter 1 was **creativity** – the ability of speakers to formulate infinite new expressions using the words and sentence patterns of the language. **Syntax** relates to the systems in place in language to support the productive creation of new sentences; this is done by combining words according to established rules to make new meanings. In linguistics, the word 'syntax' roughly corresponds with the more everyday word **grammar**. Both refer to the systems of language that support the combination and recombination of words into new sentences. We realise that 'syntax' and 'grammar' are perhaps dirty words for many people, but we hope to show you here, and elsewhere in this book, that this bad reputation is totally unjustified!

It is useful to note that there are other, more general, meanings of grammar in the community. Some associate it with the 'correct' rules of using language. As we noted in Chapter 1, although prescriptive approaches to language need to be understood and addressed within linguistics, linguists are generally more interested in what people actually *do* in using language. In the context of this book, grammar refers to our developing understanding of the structural relationships within language that allow words to be combined in meaningful ways. Some of these might be understood to be correct from a prescriptive viewpoint and some might not. This issue of language standards is addressed again in Chapter 15.

9.1 Grammaticality

> It is well to remember that grammar is common speech formulated.
>
> W. Somerset Maugham (1938), *The Summing Up*

Given our avoidance of a prescriptive stance in this chapter, it is fair to ask just how grammatical sentences will be identified. The short answer is that a grammatical sentence is one that is readily understood and that fits with the speaker's expectations about how the language is typically used. This encompasses sentences in non-standard varieties of the language, but excludes sentences that do not display normal grammatical patterns and structures. Consider the following examples.

(9.1) Savita sneezed because she'd eaten too much pepper.

(9.2) The box collapsed because it had eaten too much pepper.

(9.3) Youse be careful about all that pepper.

(9.4) Box collapse eat much pepper.

(9.5) For whom are you collapsing all those boxes?

(9.6) Box-collapsing is hazardous to your health.

(9.7) Cats like hazardous to your health.

(9.8) Jude is a black-belt in washing dishes.

(9.9) Who are you collapsing all those boxes for?

There are four possible ways to categorise these sentences.

1 **Grammatical** sentences are sentences that are easy to parse (break down into their parts) and are intuitively correct.

2 **Non-standard** sentences are grammatically correct for the variety concerned. ('Non-standard' does not mean 'substandard'. See also the discussion in Chapter 10 on variation.)

3 **Ungrammatical** sentences are not easy to parse and are intuitively wrong. Ungrammatical sentences can be indicated by

an initial asterisk: **Cat gurgle toothbrush purple.*

4 Semantically troubling sentences are grammatical and therefore can be understood, but they express propositions that do not make sense. They are ‘wrong’ in some sense – but ‘valid’ in terms of their structure. (Recall our discussion of ‘anomalous’ sentences in Chapter 6.) These sentences can be indicated by an initial question mark: ?*Our canary dictated an essay on Proust*.

How would you categorise the sentences in examples 9.1–9.9? (You can see what we think about them in the supporting materials online.)

In this chapter, we’ll continue the pattern of providing English examples as initial illustrations, but we also want to show how these concepts work in other languages from the case-study set. We often use Mali for the examples presented in the discussion; further examples from the case-study set are included in the exercises at the end of the chapter.

9.2 The basic clause

When it comes to making sentences, where do we start? A useful preliminary step for English is to begin with a basic clause. **Basic clauses** only refer to one proposition and they can stand alone as complete utterances. Basic clauses are also **positive** and **declarative**, meaning that they make a simple assertion. The pairs of sentences in Table 9.1 should help you to identify these characteristics.

Table 9.1 Contrasts between basic and non-basic clauses

Parameter	**Example**	**Complete**
Polarity		
positive	*Tia laughed.*	basic
negative	*Tia did not laugh.*	non-basic
Utterance type		
declarative	*Mish plays soccer.*	basic
interrogative	*Does Mish play soccer?*	non-basic
Structure		
simple	*Hanji bakes cakes.*	basic
complex	*Dorna likes to bake cakes.*	non-basic

A basic clause must have all three of these characteristics (positive, declarative and simple), while a non-basic clause may include any of the following combinations:

(9.10) *Why didn't Tia laugh?*[negative and interrogative]

(9.11) *Does Minh like to play soccer?*[interrogative and complex]

(9.12) *Deniz doesn't like to bake cakes.*[negative and complex]

(9.13) *Doesn't Deniz like to bake cakes?*[negative, interrogative and complex]

Any and all of these clause types can also be combined with other clauses of the same type using **coordination (**via linking words such as *and, or* and *but*).

(9.14) *Why didn't Tia laugh or even smile?*[coordinated, negative and interrogative]

(9.15) *Does Minh like to play soccer and go*[coordinated, interrogative and complex]*to the movies?*

(9.16) *Deniz doesn't like to bake cakes but*[coordinated, negative and complex]*he will make soup.*

(9.17) *Doesn't Deniz like to bake cakes or*[coordinated, negative, interrogative *was it soup that he likes to make?*and complex]

A key characteristic of all clauses is that they are made up of two parts. Characteristically, the left phrase functions as the **subject**, while the right phrase functions as the **predicate**. The basic clauses in Table 9.1, along with a few others, are divided into their subjects and predicates in Table 9.2. The subject is the main participant in the event described by the predicate.

Table 9.2 Subjects and predicates in basic clauses

Subject	**Predicate**
Tia	laughed.
Minh	plays soccer beautifully.
Deniz	bakes cakes.
The cat	is chasing a mouse.
Your purple dress	looks great with those shoes.
The trees on this side	are very shady.

As Table 9.2 shows, the predicate is the part of the clause that provides information about the subject – it attributes (or ‘predicates’) different roles, qualities and properties (what the subject is doing, what the subject is feeling, what the subject is like and so on). It will always contain a verb, and this verb may permit – or even require – other elements to complete it. These elements include such things as **objects** (*soccer, a mouse*), **complements** (*great, very shady*) and **adverbials** (*beautifully, with those shoes*). We’ll be looking closely at these different components in section 9.9, but for the moment think of objects as the entities that are acted upon by subjects, complements as the entities that assign properties or roles to subjects (and objects), and adverbials as entities that provide information such as location, manner and attitude.

The structure of sentences is sometimes illustrated using **tree diagrams**. The basic distinction between the subject and the predicate appears as in Figure 9.1.

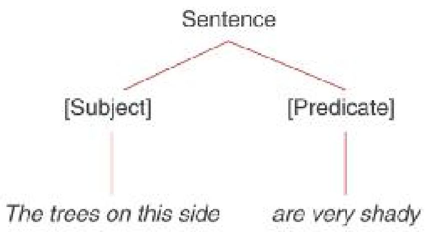


Figure 9.1 Tree diagram showing division between the subject and the predicate

Basic clauses in Mali are slightly different. They comprise a predicate with a subject pronoun. The full subject is optional and appears in topic position at the left margin of the clause. Example 9.18 shows the simplest form of the clause, containing only a subject pronoun and the verb, while example 9.19 shows a sentence containing the same verb but with additional information, including a full reference to the subject and information about how the sentence connects to the surrounding text.

(9.18)

ðlgen
be.afraid
‘I am afraid.’

(9.19)

[illegible]ki
[illegible]feminine.singular
‘But her friend was afraid.’

The hidden bias of language structure

Understanding linguistic structure is important for understanding how language interacts with aspects of human behaviour, such as memory. For instance, numerous experiments carried out by Elizabeth Loftus have shown how the simple phrasing of a question can have a substantial effect on the testimony of an eyewitness. It may even alter how the person remembers an event. One well-known piece of her research involved a hundred people who were shown a film of a multiple-car crash. Even though there was no broken glass, a question that was phrased, ‘Did you see the broken headlight?’, was more likely to produce false recognition or ‘don’t know’ responses than ‘Did you see a broken headlight?’ The definite article *the* marks a noun phrase whose referent is presumed to exist (for example, it might refer to something in the immediate context or be part of our shared general knowledge). The indefinite article *a* doesn’t make this presumption; it is used when a noun has not already been specified. The use of (in)definite articles might seem like an insignificant aspect of wording, but here it has been shown to influence people’s responses, and even their memory of an event.

In another experiment where subjects were shown different versions of a film (one containing a bicycle and one not), Loftus showed that the type of question used to interrogate subjects could influence their answers. If a person was asked, ‘You did see a bicycle, didn’t you?’ this tag question *didn’t you* was much more likely to produce a ‘yes’ response than a straightforward question framed ‘Did you see a bicycle?’, even when there wasn’t a bicycle. Tag questions are usually a request for confirmation (although this does depend on intonation); here witnesses cooperatively matched their eyewitness accounts to the expectations of the asker.

Even more worrying was how the ‘new information’ introduced through this kind of questioning then became indistinguishable from the facts acquired when the person actually did witness something – suddenly false details such as broken glass and the bicycle were incorporated into the person's memory of the events. Clearly, even a basic knowledge of language structure is going to be useful to lawyers when it comes to arguing cases!

9.3 Sentence types (simple, coordinate, complex)

Sentences are made up of one or more clauses. A simple sentence contains a single clause, while multiple clauses can either be added through coordination or embedded within the main clause by subordination. Examples of single clauses are listed in Tables 9.1 and 9.2 above.

Coordination involves the linking of two or more clauses with equal status. The most common coordinators, or coordinating conjunctions, in English are *and* (expresses addition), *but* (expresses contrast) and *or* (expresses alternatives). Consider the following examples of coordinate sentences:

(9.20) Katrina likes fruit juice *and* she likes coffee too.

(9.21) Marty likes fruit juice *but* he doesn't like coffee.

(9.22) Darla will sip coffee *or* she will eat cake.

The equal status of each of the coordinated clauses is clearly visible in a tree diagram. Note that in Figure 9.2 we have not provided details about the structure of each clause, although this would also be possible.

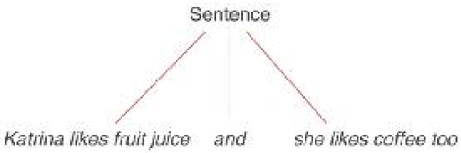


Figure 9.2 Tree diagram showing coordinated clauses

Very similar patterns can be observed in Mali. For example, in terms of coordination, Mali has forms that allow for addition (see example 9.23) and contrastive relationships (see example 9.24) and then present alternatives (see example 9.25), all marked at the beginning of the second clause. (These examples are from real texts, so they involve more complex clauses than the made-up examples used earlier, but the basic patterns still hold true.)

(9.23)

[illegible]
[illegible]male
[illegible]s-ta
[illegible]le-human. plural
‘The women spied on the boy and the men spied on the girl.’

(9.24)

[illegible]ar
[illegible]slope
[illegible]o
[illegible]n.slope
‘ … and I said to them that we should go up but they said we’d go down.’

(9.25)

[illegible]un
[illegible]pe
[illegible]ok
[illegible]t
‘(And when the war arrived), did it chase you all up higher, or did you all just stay there?’

Subordination is a little trickier than coordination; it involves embedding one clause within another, using subordinators, or subordinating conjunctions. These conjunctions also link clauses, but they don’t have the same grammatical status. Instead, one clause is subordinated to another and forms part of the main clause (e.g. in *The bread is cooked when it sounds hollow*, the beginning of the subordinate clause is signalled by *when*). Some of the typical meanings expressed by subordinators include ‘time’, ‘place’, ‘purpose’, ‘condition’, ‘reason’ and ‘concession’. Languages have many more subordinators than coordinators, and English has a number that consist of more than one word. Common subordinators in English include *after, although, as, because, if, that, until, while, provided that, as soon as* and so on.

Relative clauses are an important type of subordinate clause that is used to provide descriptions within the main clause (so they function rather like adjectives). Example 9.28 shows how the two sentences in examples 9.26 and 9.27 can be combined together using the relative pronoun *who*.

(9.26) *Darla loves Beethoven.*[simple sentence]

(9.27) *You met Darla last week.*[simple sentence]

(9.28) *Darla, who you met last week,*[complex sentence containing relative clause*loves Beethoven.*that describes subject]

Tree diagrams make the relationship between the relative clause and the main clause quite clear, with the relative clause in Figure 9.3 branching off the subject in the main clause.

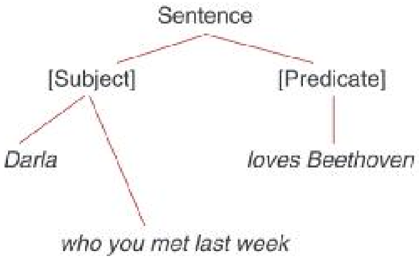


Figure 9.3 Tree diagram showing an embedded relative clause

In English, relative clauses that modify non-human referents are introduced by relative pronouns, such as *where, when, why, whose* or *which*, or by the word *that*. It is also possible to introduce a relative clause in English without any additional marking. The relative clause in each of the following examples is italicised.

(9.29) This is the place *where I had a picnic last year*.

(9.30) That was the time *when we danced till dawn*.

(9.31) Give me one good reason *why I should do that*.

(9.32) I know the lady *whose car you bumped*.

(9.33) This is the dress *which/that I prefer*.

(9.34) I love the hat *you found last week*.

Relative clauses in Mali are introduced by a special form called a **relativiser**. This form can be used with different types of referents, both human and non-human. A relativiser is always required in these constructions. Example 9.35 shows the use of the relativiser *ma* with a human noun, while example 9.36 shows its use with a non-human noun. The relative clause in Mali is in bold, while the corresponding part of the English translation is underlined.

(9.35)

[illegible]
[illegible]on- human. plural
[illegible]ngas
[illegible]
‘People who break the law are put in gaol.’

(9.36)

[illegible]
[illegible]ngular
mar-ki
relativiser big-feminine. singular
‘He dug a hole that [was] big.’

There are many other types of subordinate clauses. Another important type for English is the **complement clause**. This is a clause that fills the role of one of the main referents in the main clause. Example 9.37 shows a complement clause following the verb *try* functioning as the object. The complement clause is italicised.

(9.37) He will try *to get the coconuts down to the ground.*

The corresponding simple sentence would contain a referent – in this case, *the potatoes*.

(9.38) He will try the potatoes.

Example 9.39 is the Mali version of example 9.37. In Mali, a complement clause is often introduced by *ia*. The complement clause is bolded in the Mali translation and underlined in the English one.

(9.39)

[illegible]
[illegible]
‘He will try to get the coconuts down to the ground.’

Sentences can be both complex and coordinate.

(9.40) *You met Darla last Wednesday* [simple sentence]

(9.41) *I have known her for much longer.* [simple sentence]

(9.42) *Darla loves to play Beethoven.* [complex sentence containing a complement clause]

(9.43)

You met Darla, who loves to play Beethoven, last Wednesday, but I have known her for much longer.
[complex coordinate sentence containing relative clause (with a complement clause) that describes the subject of the first

clause]

It is clear that things can become rather complicated very quickly – clauses can be subordinated within other clauses that are themselves subordinate clauses. The following example (from the Melbourne *Big Issue* 73, p. 38) looks like a straightforward sentence, but an analysis uncovers three subordinate clauses (given in square brackets); there is also a coordinated clause linked by *and*:

(9.44) *She looked at me* [*as if I had just told her* [*I was starting up my own sect and wanted* [*to use her cat as the supreme being and her hubby for the first offering*]]].

‘A waste of energy’

Syntactic complexity is a feature of superliterate varieties such as the *-eses* (e.g. *legalese* and its offspring, *bureaucratese*). Here’s a lovely illustration in a piece of Polliespeak (from ‘A Waste of Energy’, episode 5 of the satirical TV series *The Hollowmen*). Global warming is hot on the agenda in Canberra, and the federal government is furiously back-peddling from a rather hasty announcement that threatens to derail international relations with China. This is part of the ‘official’ communiqué:

> Australia and China welcome the opportunity to affirm our commitment towards acknowledging the need to emphasise the importance of recognising the real dangers of greenhouse gas emissions.

When it comes to blurring disagreeable reality, the capability of language to pack clauses inside other clauses can provide effective camouflage. And, as many of you will recognise, the parody in *The Hollowmen* is not much of a hyperbole on the real bureaucratese it derides.

9.4 Constituency and phrases

> The linguist is of course the metaperson *par excellence*. His business is perpetual feasting on the forbidden fruit, to expose language in all its nakedness.
>
> Dwight Bolinger (1980), *Fire in a Wooden Stove: On Being Aware in Language*

The preceding two sections have focused on clauses and how clauses can be combined into sentences. Now it is time to consider how words are organised within clauses. We have already noted that in English there is a preliminary distinction between the subject and the predicate, as shown in Figure 9.1.

It is possible to expand this type of analysis, showing how the subject and the predicate themselves are structured. Units at the intermediate level of structure between individual words and clauses are called **phrases**. Figure 9.4 is an expanded tree diagram showing how words are built up into phrases and eventually combine into a whole clause that functions as a simple sentence.

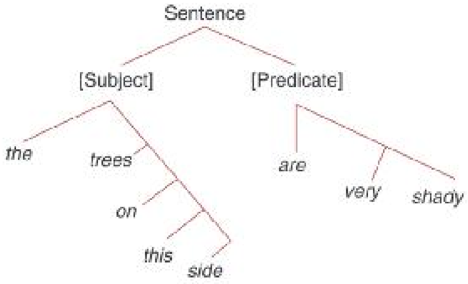


Figure 9.4 Tree diagram showing words grouping into phrases

When words are members of larger structures, such as phrases, these structures (as well as the individual words themselves) are called **constituents**. So constituents are words or groups of words that 'go together' in some way. We can say that strings of words such as *very shady, on this side* and *the trees on this side* are constituents of this sentence. A constituent can then form a close group with another constituent – that is, two constituents can together form a new constituent. If we look at Figure 9.4, the constituent *this side* combines with *on* to form another constituent, *on this side*. So, constituents exist at different levels. Think of them as linguistic Lego® pieces in the way they pattern together to form larger and larger constituent structures.

A tree diagram provides a nice overview of the structure. Each node in the tree represents a constituent and each set of branches that joins together at one of these nodes represents a phrasal constituent. Therefore, in order for any string to be considered a constituent, there must be one node that is associated with that string and with no other words.

Another strategy for showing these relationships is the use of brackets.

(9.45) [[The trees [on [this side]]] [are [very shady]]].

Brackets are handy for showing constituency (and publishers usually like them because they take up less room!). But you can see that, even in this straightforward sentence, the bracketing is complicated – and it doesn't take much for it all to become too dense and unwieldy to be helpful (and here we haven't even indicated all the constituents). Tree diagrams are a clearer way of indicating constituency (but yes, the tree is upside down – its root is at the top and its leaves are at the bottom).

As speakers of a language, we intuitively know that certain words in a sentence are linked more closely than others. Probably everyone reading this book will share intuitions that *on this side* somehow modifies the *trees* and together these words form a natural unit in this sentence – *the trees on this side*. But, how can we be so sure that words pattern in this way, other than heeding our gut feelings? In fact, we can translate these intuitions into more formal arguments, called **constituency tests**.

There are two key ways to figure out whether or not a string of words belongs together in a phrase. The first test is **substitution**: can the string of words be substituted by a single word? Within the phrase *the trees on this side*, for example, we have marked a phrase boundary between *trees* and *on this side*. It is not difficult to replace *on this side* with a single word like *there*, giving us the acceptable string *the trees there*. This is positive evidence that *on this side* is a constituent. Similarly, it is possible to replace *the trees on this side* with a single word like a pronoun, *they: **They** are very shady*.

If you are having trouble finding a single word to substitute for the string in this way, you should try to form something called an **echo question**. Basically, this involves echoing a speaker's words in order to ask for a repetition of information – or it might be a case of expressing disbelief. We substitute that part of the sentence that we want repeated with a question word such as *who, what, which, when, why* or *where*. These question words don't substitute for just any old string; the string must form a constituent.

(9.46) The trees on this side are *what*?!

(9.47) The trees *where* are very shady?

(9.48) *What* are very shady?

This test is a firm one, and if a word cannot be found to substitute for the string, this most likely means that it is not a constituent. Apply the test to the italicised strings in the following:

(9.49) They ran *up an extremely large hill.*

(9.50) They ran *up an extremely large bill.*

Superficially, these sentences look to have the same structure; however, substitution reveals immediately how different they are; we can easily find a single word to substitute for the string *up an extremely large hill*, but nothing for *up an extremely large bill.*

(9.51) They ran *there/where*.

(9.52) They ran ???

The second test is **movement**: can the string of words be moved to somewhere else? Admittedly, word order in a language like English isn't as flexible as it is in many other languages. Nonetheless, for a number of constituents, we still have at our disposal a variety of different word order arrangements that we can use for special expressive or communicative ends. In other words, a speaker may 'move' items into other positions in order, for instance, to emphasise crucial parts of their message. And the only things that can move are – you guessed it – constituents.

One important movement rule is **fronting**. Basically, this is a device whereby we shunt elements to the first position in the sentence for special emphasis or focus (*Star Wars* fans will immediately recognise this as Yoda's syntax). Let's expand our original sentence to show how this works: *The trees on this side are very shady in the hot weather*.

(9.53) *Very shady* the trees on this side are in the hot weather.

(9.54) *In the hot weather* the trees on this side are very shady.

Yoda's usual way of forming sentences is to put phrases up the front, but for ordinary speakers of English, this is a focus order that we use for special effect (we look more at this in Chapter 13). Out of context, and for certain constituents, this ordering can sound a little theatrical (but not ungrammatical).

If you try moving the suspect string to the front of the sentences (see example 9.49) and (see example 9.50), you get the following:

(9.55) Up an extremely large hill they ran.

(9.56) *Up an extremely large bill they ran.

Movement of *up an extremely large bill* results in an ungrammatical structure (hence the asterisk). Clearly these two sentences do not share the same grouping of words, and we can illustrate this with tree diagrams, as shown in Figures 9.5 and 9.6, or with brackets, as shown in examples 9.57 and 9.58. We have included brackets for only the relevant constituents so as not to muddy the waters. We have also used triangles in the tree diagrams – a handy short-cut for showing one level of constituents but not another. In this way, we can spare you lower levels of relationships within the phrases *an extremely large hill* and *an extremely large bill.*

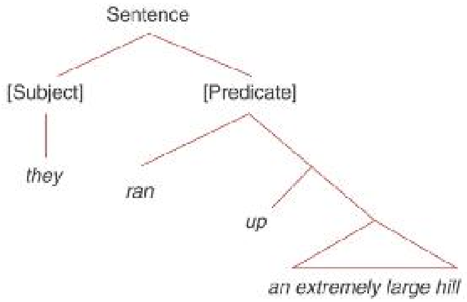


Figure 9.5 Tree diagram showing the constituency of *they ran up an extremely large hill*

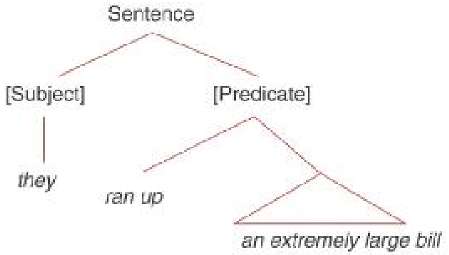


Figure 9.6 Tree diagram showing the constituency of *they ran up an extremely large bill*

(9.57) [They] [ran] [*up an extremely large hill*].

(9.58) [They] [ran *up*] [*an extremely large bill*].

You probably feel intuitively that the string *ran up* forms a natural unit in example 9.58 but not in example 9.57, and the constituency tests will back you up here.

To sum up, clauses are the principal units of which a sentence is made. Between the clause and the word, we have intermediate constituents known as phrases. Our writing system provides only a very inadequate and sporadic representation of these intervening units, but basically these are groups of words that have some sort of grammatical relationship with one another. Phrases can actually form a close group with another phrasal constituent; that is, two phrases can together form a new phrase. Looking again at the earlier sentence, the phrase *very shady,* when combined with *is* forms another phrase *is very shady*. Constituents, then, are like building blocks – they combine together to form ever larger units of structure. This hierarchical structure is characteristic of all human languages: Sentence – Clause –

Phrase – Word – Morpheme.

9.5 Categories and functions

In thinking about how clauses and sentences are structured, there is one final distinction that is helpful. This is the difference between the **category** of a word, phrase or clause, and its **function**.

- Category labels identify what kinds of elements we are discussing.
- Function labels describe relationships between different parts of a clause.

Many of the relevant category labels for words were introduced in the chapter on morphology (see section 5.6), when we looked at word classes and parts of speech such as nouns, verbs, adjectives and adverbs. All the four major word classes of nouns, verbs, adjectives and adverbs have corresponding phrasal categories – namely, noun phrases, verb phrases, adjective phrases and adverb phrases. These phrases take their name from their most important constituent word, called the **head**. For example, a noun phrase must always contain some sort of head noun and is not grammatical without one. If we take the noun phrase *those four sinister-looking limousines,* the word *limousines* here is the head. It is the core part of the phrase and all the other components are somehow dependent on it. Not surprisingly, these are called **dependents**. They are subordinate elements and they generally have a sort of modifying or describing function. Unlike the noun *limousines,* elements like *those, four* and *sinister-looking* are completely optional. We can omit them quite happily without affecting the grammaticality of a sentence, as in *limousines arrived.*

Noun phrase thrombosis

One tricky aspect of English has always been its potential for complexity in its noun phrases. Noun phrases can be linked to other noun phrases and their modifiers can themselves be modified (the image is one of ever-nesting constituents – think of Russian dolls or Chinese boxes). A bread recipe might simply recommend that *we can achieve elasticity by kneading*, or it might tell us:

> The induction of a soft elasticity into the bread dough and the promotion of a smooth and even texture in the end product can be achieved by the operation of enthusiastic kneading and in the initial stages of the procedure the addition of subtle flavoured oils together with sugar for assistance in the conversion of the flour’s natural sugars into carbon monoxide.

This example illustrates the horrifying potential of this particular aspect of noun-phrase modification (and those brave ones among you could attempt a tree diagram or bracketing to show the constituency of noun phrase heads and modifiers). Of course, none of us speaks like this, but such examples make regular appearances in some written varieties of English, such as the *-eses* (e,g, *legalese* and *bureaucratese*) described earlier.

In Figure 9.7, we have expanded Figure 9.4 to show both word- and phrase-level labels. The substitution test shows us that we have intermediate structure here; for example, we can show that *the trees on this side* is a unit because we can substitute it with single words such as *those* and *ones* (‘Those/ones in the shade’). But, we don’t want to go into what we call these intermediate phrases: we leave that to a subject devoted to syntax!

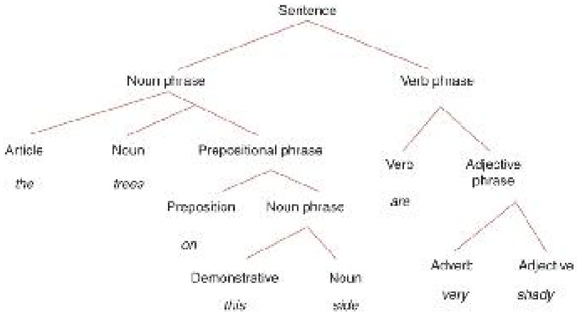


Figure 9.7 Tree diagram showing labelled words and phrases

Functions describe the role of particular elements within the clause. In short, phrasal categories like noun phrases and verb phrases are what constituents **are**, and functions like subject and object are what constituents **do**. So, the labels ‘subject’ and ‘predicate’ are function labels, since they describe the different functional roles of these two constituents.

Constituents within phrases also have specific functions. For example, in addition to a head noun, noun phrases often contain a determiner such as an article (*the, a*), a demonstrative (*that, those*) or a possessive noun phrase (*her, John's*), and possibly additional modifiers of some sort (often an adjective phrase).

The ability of heads of phrases to stand alone means that phrases may consist of only one word – hence *limousines* in the earlier sentence *limousines arrived* is both a noun and a noun phrase. You might think this a little odd, but it is grammatically justified. All the rules that apply to heads will also apply to their phrases – hence, these functions may be associated with single words or with larger phrases. For example, the simple possessive noun phrase *Sarah's* in Figure 9.8 could be expanded into something longer, such as *my best friend Sarah's*. In Figure 9.9, we have included a modifier, a prepositional phrase *on the top shelf*, this time following the head noun *book*.

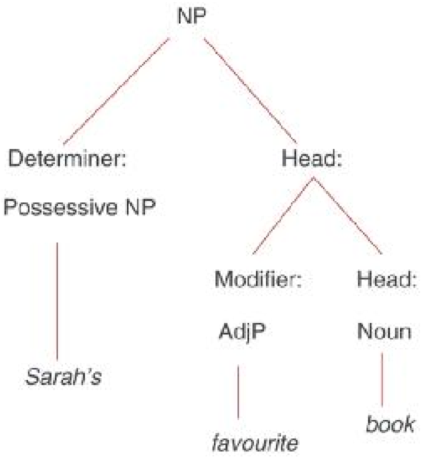


Figure 9.8 Tree diagram showing functional categories in the noun phrase *Sarah's favourite book*

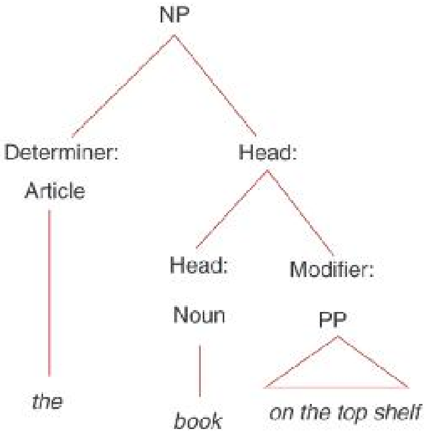


Figure 9.9 Tree diagram showing functional categories in the noun phrase *the book on the top shelf*

Syntax and the law

When people write, they often leave behind clues about who they are. Many linguistic features, including those of syntax, can be idiosyncratic and can help determine authorship when this is in dispute. This is where linguists come in.

Robert Eagleson was one such linguist called in to provide expert testimony to help identify the true author of a letter. A Sydney woman had been missing for some time and her husband was suspected of murdering her. He then produced a six-page letter that he claimed she had written as a farewell to her children. Because there was no trace of the woman's body, the identity of the author became crucial. Eagleson compared the disputed letter with writings from both the husband and wife, but since the letter was typed, it was not a matter of assessing actual handwriting. Spelling errors became critical – so did aspects of morphology and syntax. It turned out that the farewell letter and the husband's writing shared a number of these features: there were too many and they were too consistent in nature to be coincidental. Both were missing present tense and past tense suffixes (for example, *he threaten me; she had a knife and threaten during argument*). Both shared a sentence structure whereby clauses were simply juxtaposed and not clearly demarcated (*since his accident at work he's slowed down before that he wanted it everynight*). Crucial grammatical words, such as prepositions, were systematically omitted (for instance, missing *of* from *hundreds … dollars*; missing *to* from *whether … start*).

Similarities of linguistic practice between the farewell letter and the husband's writing, as well as the lack of comparability between the letter and the wife's writings, suggested that the husband was in all likelihood the author. Even though he had pleaded his innocence and had continued to assert that his wife had written the letter, the linguistic evidence was overwhelming. After the case went to trial, he changed his plea and admitted to killing his wife and writing the farewell letter.

9.6 Grammatical relations and semantic roles

One of the basic tasks of language is to take situations in the real world, and map them onto the rules and structures of the language in order to express appropriate and relevant meanings. Languages vary in the aspects of a situation they encode in detail and also in the extent to which different versions of the same event can be encoded. The passive construction in English is useful for showing how different **semantic roles** (played by the different constituents) can be mapped to **grammatical relations** (the functional relationships between these constituents). Consider the following examples that all describe the same event but portray it in different ways:

(9.59) My grandfather built that house back in 1952.
(9.60) That house was built back in 1952.
(9.61) That house was built by my grandfather.

If we focus on the noun phrases *my grandfather* and *that house*, we can see that they can both be used in subject function in different representations of the situation. Furthermore, the noun phrase *my grandfather* can be completely removed from the sentence in the version in example 9.60. What is going on here?

One way to think about these different sentences is to draw out a distinction between the semantic roles that are involved in an event and the grammatical functions that they are assigned to. In both example 9.59 and example 9.61, the noun phrase *my grandfather* refers to the semantic role of the agent – the person who built the house. However, whereas in example 9.59 the noun phrase *my grandfather* is in subject position, in example 9.61 this noun phrase is introduced using the preposition *by* and occurs at the right margin of the sentence. Example 9.60 shows that the prepositional phrase *by my grandfather* is optional. In all three sentences, the noun phrase *that house* is semantically a patient in this event. It is acted upon, brought into being, by the agent. In example 9.59, *that house* is treated as the object of the verb, but in example 9.61, it is treated as the subject. In other words, the syntax of the three sentences is different, but the semantic roles remain the same.

In English, this pattern of removing the agent subject and promoting the object to subject position is called **passivisation**. It also involves a special form of the verb, taking the auxiliary verb *be* and the past participle form of the main verb. We will return to this process and why we use it when we look at information structure in Chapter 13. For the moment, though, consider Figure 9.10, which illustrates how the grammatical relations work out when we convert a simple active sentence like *My grandfather built that house* into the equivalent passive sentence *That house was built (by my grandfather)*.

Many languages have a grammatical structure very similar to the passive in English (although it may be used for different purposes). In addition, there are some languages – and Japanese is one of these – that also have something called the adversative passive. In its structure, this works like the ordinary passive except it signals that the subject is unfavorably or unpleasantly affected by the action or the event described by the verb.

In terms of semantic roles, some common types of participants include agent, experiencer, theme, patient, possession, instrument, location, goal, source, path, recipient and beneficiary. Languages vary in how these different roles are encoded. In languages that make use of a lot of case marking, these roles are handled in connection with case assignment. Consider the relationship between semantic roles and cases in relation to Russian, as shown in Table 9.3.

Table 9.3 Russian cases mapped to semantic roles, with English examples

Russian case	**Thematic roles**	**English example**
Nominative	agent	*My sister* made the cake.
	experiencer	*Gregor* heard the song.
Accusative	patient	Anna spilt *the milk*.
	theme	My aunt loves *cats*.
Genitive	possession	*Sarah's* book
Dative	recipient	He gave *his mother* some flowers.
	beneficiary	I made this cake *for you*.
Instrumental	instrument	They dug the garden *with spades*.
	theme (with body parts)	She waved *her hand*.

Prepositions (or postpositions) are also important means of encoding this information. Russian expresses other semantic roles with prepositions. Languages may therefore make use of multiple strategies in assigning these roles.

Other languages rely on word order to establish the main relationships. English uses word order to distinguish between subjects and objects, and normally assigns agents and experiencers to the subject role, and themes and patients to the object role. Having no case system also means that English is quite flexible as to what semantic roles can be mapped onto these grammatical relations. The following examples give some idea of the range that can appear as subject. For comparison, we have provided the (ungrammatical) equivalents in German, a close relative to English. As in Russian, the German case system means the language is more restricted in the semantic roles that can appear as subjects and objects.

(9.62) This tent sleeps only four people.[= location]
***Dieses Zelt** schläft nur vier Leute.

(9.63) The red wine drinks well. (Waitspeak!)[= patient]
***Der Rotwein** drinkt gut.

(9.64) The latest edition of the cookbook has added thirty recipes.[= recipient]
***Die letzte Ausgabe des Kochbuches** fügte dreißig Rezepten hinzu.

(9.65) Tomorrow promises to be cold and wet in most places. [= time]
***Morgen** verspricht meistenorts kalt un naß zu geben.

German subjects are normally agents (in the nominative case), and other roles are rendered by cases (other than the nominative) or by prepositional phrases.

Many semantic roles in English are also encoded using prepositional phrases.

(9.66) Only four people can sleep *in this tent.*[= location]
Thirty recipes have been added *to the latest edition of the cookbook.*[= recipient]

More on that ‘awful German language’

In Chapter 5, we introduced you to Mark Twain’s struggle with the German language, a struggle that he described in several different works and speeches. He even wrote an essay (published in 1880) entitled ‘The Awful German Language’, in which he outlined the terrors that confronted speakers of English having to learn German. Since German is one of our case-study languages, we feel we can give you another taste of his satirical account of its grammar; here he is tackling the thorny issues of cases, prepositions and semantic roles.

> [M]y book inquires after a certain bird – (it is always inquiring after things which are of no sort of consequence to anybody): ‘Where is the

> bird?' Now the answer to this question – according to the book – is that the bird is waiting in the blacksmith shop on account of the rain. Of course no bird would do that, but then you must stick to the book. Very well, I begin to cipher out the German for that answer. I begin at the wrong end, necessarily, for that is the German idea. I say to myself, 'REGEN (rain) is masculine – or maybe it is feminine – or possibly neuter – it is too much trouble to look now. Therefore, it is either DER (the) Regen, or DIE (the) Regen, or DAS (the) Regen, according to which gender it may turn out to be when I look. In the interest of science, I will cipher it out on the hypothesis that it is masculine. Very well – then THE rain is DER Regen, if it is simply in the quiescent state of being MENTIONED, without enlargement or discussion – Nominative case; but if this rain is lying around, in a kind of a general way on the ground, it is then definitely located, it is DOING SOMETHING – that is, RESTING (which is one of the German grammar's ideas of doing something), and this throws the rain into the Dative case, and makes it DEM Regen. However, this rain is not resting, but is doing something ACTIVELY, – it is falling – to interfere with the bird, likely – and this indicates MOVEMENT, which has the effect of sliding it into the Accusative case and changing DEM Regen into DEN Regen.' Having completed the grammatical horoscope of this matter, I answer up confidently and state in German that the bird is staying in the blacksmith shop 'wegen (on account of) DEN Regen.' Then the teacher lets me softly down with the remark that whenever the word 'wegen' drops into a sentence, it ALWAYS throws that subject into the GENITIVE case, regardless of consequences – and that therefore this bird stayed in the blacksmith shop 'wegen DES Regens.'

Syntactic ambiguity

Speakers recognise that a sentence can hold more than one meaning and linguists have to be able to deal with this. For example, the following sentence is ambiguous in two ways:

(9.67) The druids decided on the village green.

This is an example of structural ambiguity that we can capture very simply with brackets to show the constituency:

(9.68) The druids [decided on] [the village green].

(9.69) The druids [decided] [on the village green].

Sentences like this one, with more than one syntactic analysis, can be captured with paraphrase (that is, we can restate them using other words):

(9.70) The druids came to a definite conclusion about the village green.

(9.71) The druids made up their minds while they were on the village green.

Part of speakers' competence in a language is that they can recognise when sentences mean the same thing – in other words, their relationship is one of paraphrase. For example:

(9.72) The dog bit the cat – the cat was bitten by the dog

(9.73) I gave the druids the book – I gave the book to the druids

(9.74) I ate up my quiche – I ate my quiche up

(9.75) Teaching syntax is fun – It's fun to teach syntax

Linguists have to be able to capture this competence in an analytical (and rigorous) way, and in the Further reading section at the end of the chapter, we suggest where you can follow this up.

Following prominent speakers over the edge

Some time back, a statement appeared in the *Advertising and Community Services Bulletin* of a university known to us (which shall remain nameless). It read, 'On Wednesday, [we] will participate in an International Day of Student Action. At 1pm, a banner will be dropped from the building, followed by three prominent speakers.' Here is one fine example of unfortunate ambiguity. A phrase like *followed by three prominent speakers* is a mini-clause containing the participle *followed*. Such a phrase should be attached to something, but in this example it is left dangling, and this produces the hilarious image of a day of action that potentially involves dropping a banner from the top of the building and then dropping three prominent speakers.

Linguistic misadventures like this one are especially obvious in writing. In speech, we get away with of all sorts of linguistic near-misses that are never picked up by our audience. There is too much happening, and it's all happening far too fast for them to notice the infelicities of style and structure that go whizzing by – besides, there are all sorts of other cues (intonation, pitch, pausing and so on) to help overcome potential problems. Writing, however, is a different kettle of fish. Dangling constituents like the one above have a comical effect and are therefore distracting to readers. If something is distracting, then it will interfere with effective communication and even laid-back linguists, like us, would say that in writing, that is best avoided.

9.7 Phrase structure rules

Recall that all the English sentences we have examined so far have been composed of two major constituents: a noun phrase (subject) followed by a verb phrase (predicate). This generalisation can be expressed in the form of something called a **phrase structure rule**:

(9.76) S → NP VP

In other words: S(entence) → N(oun)P(hrase) V(erb)P(hrase). This rule states that the constituents of a sentence (S) are a noun phrase (NP) followed by a verb phrase (VP); the arrow means something like 'consists of' or 'rewrites as'. So, this represents another way of presenting a syntactic analysis.

If you look back at our constituent structure tree for *The trees on this side are very shady*, you will see that there are other rules that reflect generalisations about the syntactic structure of English sentences. A noun phrase, for example, can consist of a noun, a determiner and a noun (DET N), and potentially also a preposition phrase (PP) modifier.

(9.77) *Trees* are very shady.(N)

(9.78) *The trees* are very shady.(DET N)

(9.79) *Trees on this side* are very shady.(N PP)

(9.80) *The trees on this side* are very shady.(DET N PP)

We can express all this information in the following neat way (the parentheses indicate that the determiner and the prepositional phrases are optional):

NP → (DET) N (PP)

There are also all sorts of possibilities for verb phrases; they can contain adjective phrases (AjP), noun phrases and prepositional phrases:

(9.81) The trees on this side *flourish*.(V)

(9.82) The trees on this side *are very shady*.(V AjP)

(9.83) The trees on this side *are very shady in summer*.(V AjP PP)

(9.84) The trees on this side *give really good shade*.(V NP)

(9.85) The trees on this side *give really good shade in summer*. (V NP PP)

We can capture these possibilities by the following phrase structure rule:

(9.86) VP →V ({AjP NP}) (PP)

Here you can see that, in addition to parentheses, we have introduced another shorthand notation: the curly bracket. It means something like ‘pick one from’ – in other words, only one of the items in the brackets can be chosen (we can’t have an adjective phrase and a noun phrase appearing at the same time). Adjective phrases can themselves take modifiers, such as the adverb (Av) *very* in (9.83); we can also capture this with parentheses:

(9.87) AjP → (Av) AjP

And the structure of prepositional phrases can be described as:

(9.88) PP → P NP

Inspired by the work of Noam Chomsky, this type of grammar is called a **generative grammar** (sentences are ‘generated’ by the operation of these rules). The power of these rules is that this small set can generate the constituent structure tree that was given for *The trees on this side are very shady*, as well as a large number of other sentences, too. Of course, just these few rules don’t begin to account for all the possible sentence structures found in English, but, hopefully, you can see the potential of this kind of grammar.

One problem with our little set of rules here is that they will also generate semantically ill-formed sentences, such as our earlier *the box ate too much pepper* or Noam Chomsky’s famous *colourless green ideas sleep furiously* – in other words, sentences that are grammatically well structured but unacceptable because they are anomalous. To rule out such sentences, we would require another set of rules (known as **lexical insertion rules**), which would specify which words could be used. Here’s an example of what they might look like for this small set of phrase structure rules:

(9.89)

N → {tree, shade, summer}

DET → {the, this, a}

V → {flourish, are, give}

Aj → {shady, good}

P → {on, in}

Av → {very, really}

Of course, we can always assign figurative meanings to the anomalous items in those semantically dodgy sentences: *box* might refer to a friend built like a container; *green* and *colourless* are perfectly compatible if we ignore their literal senses and assume they are metaphors (*green* as ‘immature’ and *colourless* as ‘uninteresting, dull’); and so on. So, it is controversial whether or not you want your grammar to generate such sentences.

Phrase structure rules have been developed for many other languages. Consider the following example sentences from Japanese (many papers have been written around these sentences, making Ken’s love for Naomi famous in a very small circle of people).

(9.90)

Ken *ga* ***Naomi*** *wo* ***aisiteiru***
Ken subject **Naomi** object
‘(We’re talking about Ken and I am telling you …) Ken loves Naomi’

(9.91)

Naomi *wo* ***Ken*** *ga* ***aisiteiru***
Naomi object **Ken** subject
‘(We both know Naomi and I am now telling you that …) Ken loves Naomi.’

There are three notable differences regarding the English examples we have been using so far. First, the verb occupies the final position in the sentence (as opposed to the position after the subject and before the object, if there is one). Second, nouns occur in phrases marked by postpositions (in this case, *ga* and *wo*), indicating the function of the noun phrase. Finally, word order in Japanese is also more variable than in English, so the rules need to be able to account for different ordering options (and allow Naomi to reciprocate if she wishes). Here is a set of rules that allow us to formulate either order of noun phrases. Note that in example 9.93, TV stands for ‘transitive verb’. (Note that *ga* and *wo* are particles and are sometimes represented as affixes rather than independent words.)

(9.92)

S → PP VP

VP → PP TV

PP → NP P

NP → N

Again, lexical insertion rules are required to complete the process.

(9.93)

TV → *aisiteiru*

N → {*Ken, Naomi*}

P → {*ga, wo*}

Recursion

One important aspect of language that these sorts of rules can handle well is creativity. This is something we have mentioned at several points already in this book: words can combine and recombine in an infinite number of structures. Natural languages have the property that they allow infinite **recursion** (or iteration) of categories. A simple example is the ability of speakers to put sentences inside other sentences.

(9.94) Mary believes Bill knows Fred

(9.95) Mary believes Bill knows Fred knows John

(9.96) I hope Mary believes Bill knows Fred knows John …

There is no theoretical limit to the length of these types of sentences, and the strength of these phrase structure rules is that they allow recursive rules that handle this feature easily. The sort of rule that would allow for sentences like the ones above would look like this:

(9.97) VP → V S

Note that we can also collapse our two-verb phrase (VP) rules using the curly brackets again – in other words, VP can expand as either V NP (verb + noun phrase) or V S (verb + sentence).

(9.98) VP → V {NP, S}

So, the phrase structure rule generates a VP, and the VP rule generates a sentence (S), which then generates a VP, and so on! Because of this recursive aspect, a very small set of rules can generate a wide range (in fact an infinite number) of sentences.

Another aspect of recursion is that certain phrases (such as adjective phrases) can be repeated, and there is theoretically no limit to the number of modifiers – there is a practical limit, though (speakers will run out of puff, and audiences most certainly will run out of patience!). The following advertising caption for Nuttall's Mintoes is a wonderful example of repeated adjective phrases:

> Why do you think we make Nuttall's Mintoes such a devilishly smooth cool creamy minty chewy round slow velvety fresh clean solid buttery taste?

The house that Jack built

Good examples of recursion are 'the house that Jack built' structures that you will know from nursery rhymes like the following:

> This is the rooster that crowed in the morn
> That woke the judge all shaven and shorn
> That married the man all tattered and torn

That kissed the maiden all forlorn
That milked the cow with the crumpled horn
That tossed the dog that worried the cat
That chased the rat that ate the cheese
That lay in the house that Jack built.

The cumulative tale builds on the ability of noun phrases (NPs) (such as *the rooster*) to have relative clauses (such as *that crowed in the morn*) that themselves have NPs able to be modified (the noun *morn* has the relative clause *That woke the judge all shaven and shorn*) – and so it goes on, just like 'the house that Jack built'. All we need is to do is to change our NP rule so that it can have a sentence (S) (in this case, a relative clause). And because NPs are themselves constituents of S, we have a rule that recursively feeds our expansion of S, again allowing for sentences that are infinitely long.

9.8 Word order types

> Whenever the literary German dives into a sentence, that is the last you are going to see of him till he emerges on the other side of his Atlantic with his verb in his mouth.
>
> Mark Twain (1889), *A Connecticut Yankee in King Arthur's Court*

Because all languages of the world show a range of similarities and differences in their structure, it is possible to set up language types – that is, groups of languages classified according to the features that they have in common and that differentiate them. We can base **typological groupings** of languages on whatever features we want to, and in Chapter 11 we will see an important typology based on morphological structure.

Here we consider another central typological classification based on the order of major constituents: S(ubject), V(erb), O(bject). Of all the mathematically possible word order combinations across these different phrasal categories, languages utilise surprisingly few; the majority fall into SVO, SOV or VSO. These typological classifications may or may not coincide with genetic classifications (based on shared inherited features). Unrelated English and Mali belong to the same typological group in that they share the basic word order: S(ubject), V(erb), O(bject). As we saw in section 9.7, Japanese is verb-final (SOV) and shows some freedom about how S and O are ordered. German and English are relatives, but they differ typologically;

German has split word order – SVO(V) in main clauses (complex verbs are split and some parts go to the end) and SOV in subordinate clauses (and in literary works, verbs can appear at the end of some very long sentences, as suggested by Mark Twain's image at the start of this section!). Russian is also related to English, but it has very flexible word order – all possible permutations of S, V and O are grammatical (although SVO is more usual). So, languages regularly belong to one genetic group, but to different typological groups.

The order VSO, with the verb in first position, is less usual than the other two. Only one of our case-study languages, Māori, has this order. As the following examples show, tense markers come first, followed by the verb, the subject (here bolded) and then everything else.

(9.99)

[illegible]
[illegible]
'The meeting is over'

(9.100)

[illegible]
[illegible] snatch
'I snatched at the gun'

(9.101)

[illegible]
[illegible] heavy
'The beat of the rain was heavy'.

(9.102)

[illegible]
[illegible]
'I feared that he might return home.'

9.9 Transitivity

The verb is very important in the process of managing the relationship between semantic roles and grammatical relations. Each verb characterises an event in a particular way, and this determines how many participants must be represented in an event. These relationships between the event and its participants, encoded by the verb, are called different **transitivity** patterns. (The term *transitive* is associated with the idea that there is a transfer of action between the agent and the patient in an event). There are five basic transitivity patterns in English, depending on what elements are needed in order to form a grammatical sentence.

Intransitive clauses have one participant in subject position.

(9.103)

The flowers	*blossomed.*
Subject	

Transitive clauses have two participants: a subject and an object.

(9.104)

My cat	*shredded*	*my black stockings.*
Subject		Object

Decision in law comes down to a question of transitivity

> ‘I know it was completely wrong, but I thought fuck it, I might as well fuck it.’

These were the recorded words of one of the top Westpac Bank traders Colin Roden. It was all about rigging interest rates, and the whole case hinged on the transitivity of the verb *fuck* in Roden’s words. Was it transitive and therefore doing something to an object (so fitting the pattern X verbed Y) – the meaning here would be ‘destroy, ruin’, and the following object *it* would be a key benchmark interest rate? Alternatively, was this use of *fuck* intransitive – with more an expressive meaning that captured the ‘I-couldn’t-care-less’ ethos of the time (the *it* in this construction would be empty, the sort of dummy object you find in other expressions like *cark it* and *leg it*)? Unfortunately for Westpac, the Honourable Justice Jonathan Barry Rashleigh Beach of the Federal Court ruled in favour of transitivity; in other words, the judge found Mr Roden had planned a ‘doable activity’.

Ditransitive clauses have three participants: a subject, an indirect object and a direct object.

(9.105)

My [illegible] *shelf.*
Subject [illegible] object

Complex intransitive clauses have a subject that is described by a predicative complement (these are explained in more detail below).

(9.106)

Your little sister [illegible]
Subject Predicative complement

Complex transitive clauses have a subject and an object that is described by a predicative complement.

(9.107)

[illegible] *ncer.*
Subject Object Predicative complement

The function of subject does not play any role in the classification of verbs, since all verbs in English need a subject. Important for this classification are the functions of object, direct and indirect object, and predicative complement. These are required by particular types of verbs, and for this reason they are called **complements**. Note that it is important not to confuse this term with the function label 'predicative complement'. Complement here is a general term; it refers to any element that is required by a head, be that a verb, a noun, an adjective or a preposition.

Earlier, we touched briefly on another function: adverbial. These are the more peripheral elements of the clause, and are usually optional. For example, adverbials like *always* or *occasionally* can freely be added (or not) to all the clauses above without disturbing their grammaticality. By comparison, these clauses will become unacceptable or their meaning will change significantly if an object or predicative complement is omitted. In this next section, we describe a few key identifying features of complements and adverbials in English. However, before doing so, we address some of the structural characteristics of subjects.

The subject

You will find that many textbooks use semantic rather than grammatical criteria to define the subject. For example, they will point out (as we've done here) that subjects often involve the agent (sometimes called the actor). This means that subjects are the 'doers'. Certainly it is the case that subjects are often the instigators of an action expressed in the verb string but, as Table 9.4 on p. 287 shows, English subjects can have all sorts of roles – including patient (is affected by another participant), instrument (is involved in an action but controlled by another participant), location (is situated in space), experiencer (undergoes a physical effect from some stimulus), recipient (receives something), and so on. In something like *it is raining,* the subject *it* doesn't appear to have a semantic role at all. Functions like subject are always best defined in terms of clusters of structural features. For English, these include the following.

Table 9.4 Relics of cases in the English personal pronoun paradigm

	Nominative		**Accusative/dative**	
	SG	**PL**	**SG**	**PL**
1	*I*	*we*	*me*	*us*
2	*you*	*you*	*you*	*you*
3	*he, she, it*	*they*	*him, her, it*	*them*

Phrasal category

Subjects are either noun phrases (including single nouns and pronouns) or subordinate clauses. Examples 9.108 and 9.109 show two different kinds of categories acting as subjects. The first is a noun phrase, while the second is a subordinate clause (the tree diagram in Figure 9.11 illustrates this).

(9.108)

[Subject noun phrase] *Minh's mother* is a sharp chess player.

(9.109)

[Subject subordinate clause] *To say Minh's mother is a sharp chess player* is a complete exaggeration.

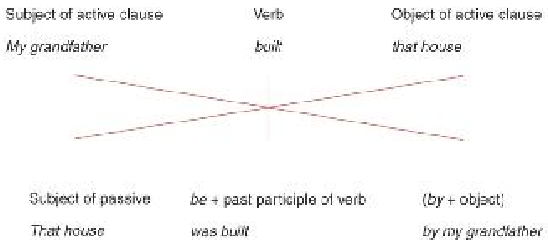


Figure 9.10 Diagram showing active and passive versions

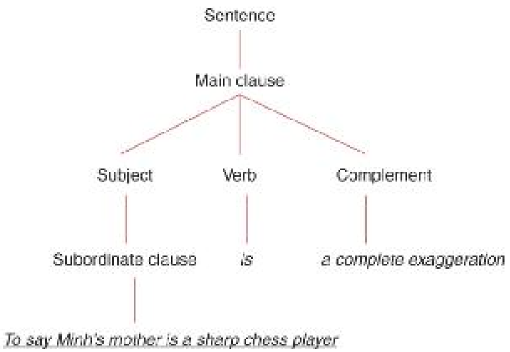


Figure 9.11 Figure showing a complex sentence with a subordinate clause subject

Basic position

The subject always precedes the predicate in basic clauses (example 9.110). However, in interrogatives, the subject appears between the auxiliary and (any) main verb (example 9.111).

(9.110) *Tia* ate the pumpkin soup.

(9.111) Can *Sofi* really do that?

Pronoun form

English doesn't really have case anymore, but some pronouns have retained distinctive case forms. (Recall that in Old English, case was marked on many elements within the noun phrase and was an important indicator of the structure of the clause, but the modern language relies on word order to a much greater extent). Table 9.4 shows the relics of old nominative and accusative/dative cases in the Modern English personal pronoun paradigm, used for subjects and objects respectively.

Verb agreement

The subject noun phrase determines the form of the verb (although this criterion is rather limited, since verbs only have different forms in the present tense). Note that if we change the person and number of the subject in the following examples, the form of the verb then has to change.

(9.112) I coach my son's soccer team.

(9.113) They coach their son's soccer team.

(9.114) He coaches his son's soccer team.

Obligatoriness

In English, it is necessary to have a subject in basic clauses, and if we removed the subject in all the previous examples, these sentences would be ungrammatical. However, in casual speech and in some written varieties, such as e-speak, this requirement is relaxed (e.g. *am running a bit late*).

The object

There are two types of objects: direct objects and indirect objects. Although most of what we describe here holds for both types of object, they do differ semantically. The indirect object is one function where a semantic definition actually works quite well. As in the following example, it typically involves the semantic role of either recipient or benefactor (or malefactor in situations where the entity is adversely affected), meaning that it is the participant who receives something or benefits from the action. Accordingly, verbs that can take indirect objects are verbs of giving, wishing, communicating, causing and so on.

(9.115)

[illegible] *parents.*
[illegible] object

In many cases, it is possible to present the same situation using a single object (equivalent to the direct object) and a prepositional phrase (encoding the same participants as the indirect object). Recipients are introduced by the preposition *to* and beneficiaries are introduced by the preposition *for*.

(9.116)

[illegible] *grandmother.*
[illegible] Prepositional phrase

(9.117)

[illegible] *yone.*
[illegible] Prepositional phrase

The semantic role of direct objects is more complex; they traditionally have been linked to the noun phrase that is most affected by the action of the verb. The direct object is frequently described as the participant that has the role of patient, the participant that is affected by the action. In actual usage, though, objects cover a much wider range of semantic roles than simply the ‘patient’ or ‘target’ of the action.

In basic clauses, objects appear directly after the verb and, like subjects, they are either noun phrases (including single nouns and pronouns) or subordinate clauses (see Figure 9.12). As Table 9.4 indicated earlier, these pronouns have a distinct object (accusative) case form (*them* versus *they* in example 9.119).

(9.118) We are cooking *chips*.

(9.119) You drain *them* next.

(9.120) Most outlets advertise *that they use cholesterol-free oil*.

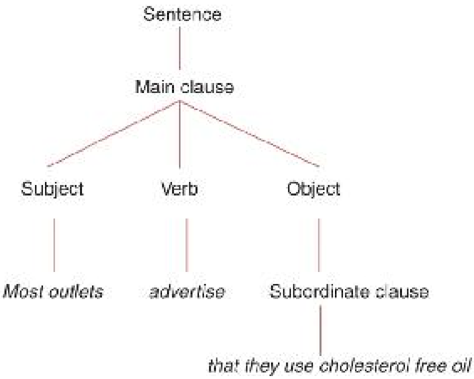


Figure 9.12 Figure showing a complex sentence with a subordinate clause object

Predicative complements

Like objects, **predicative complements** occur after the verb. How can we tell them apart? Predicative complements have the function of providing additional information about another entity in the clause – they do not refer to a new entity. Some English verbs can take either an object or a predicative complement. Examples 9.121 and 9.122, using these types of verbs, help to show the difference.

(9.121)

[illegible]m.

(9.122)

[illegible]

[illegible] complement

It is also possible for a predicative complement to refer to an object (although this use is limited in some varieties of English).

(9.123)

[illegible] *entertaining.*

[illegible] complement

(9.124)

[illegible] *friend*

[illegible] complement

Adverbials

As described earlier, adverbials are typically the non-essential elements in a clause; usually they can be added or removed freely from any clause. So, while clauses permit at most one subject, one finite verb, one predicative complement, one direct object and one indirect object, there are no such restrictions with respect to the number of adverbials. In fact, there is no theoretical limit to the number of adverbials we can include – in principle, we can keep adding them indefinitely. Typically, adverbials describe when, where or how something happened, so we can then have a criterion that an adverbial is the answer to a question beginning with something like: *When? Where? Why? How? How long? How often? or With whom?* but they never answer questions like *what?* or *who?* Table 9.5 presents examples of some of the different types of adverbials found in English; as you can see, they can appear different types of phrases and even clauses.

Table 9.5 Semantic types of adverbials

Category	**Type**	**Example**
Manner	Adverb phrase	Deniz and Jackson watched the movie *reluctantly*.
Place	Prepositional phrase	Deniz and Jackson watched the movie *at the cinema*.
Time	Noun phrase	Deniz and Jackson watched the movie *last week*.
Duration	Prepositional phrase	Deniz and Jackson watched the movie *for ten minutes*.
Frequency	Noun phrase	Deniz and Jackson watched the movie *every week*.
Purpose	Subordinate clause	Deniz and Jackson watched the movie *to find out what happened*.
Condition	Subordinate clause	Deniz and Jackson will watch the movie *if they must*.
Concession	Subordinate clause	*Although they were reluctant*, Deniz and Jackson watched the movie.

The other characteristic of adverbials is that they are quite flexible in their positioning. Unlike subjects and objects, they move about the sentence with ease. Depending on emphasis, there are a number of different versions of each of the sentences in Table 9.5 – for example, *Deniz and Jackson watched the movie reluctantly; Reluctantly Deniz and Jackson watched the movie; Deniz and Jackson reluctantly watched the movie.*

The 'great Australian adjective' that isn't

Since the 1880s, the word *bloody* has been described as 'the great Australian adjective'; however, we feel duty bound to point out that it's not actually an adjective – and here's why.

There are three things that characterise adjectives. Contrast something like *shut that rusty gate* with *shut that bloody gate.* All well-behaved adjectives can appear immediately before the noun they are describing. So, a tick here for both the words *rusty* and *bloody*. However, well-behaved adjectives should also be able follow verbs like *be* or *become* (in other words, as subject predicates). So, we can alternate between *that rusty gate* and *that gate is rusty*. But, notice that *that gate is bloody* is not a paraphrase for *that bloody gate.* Here *bloody* – the so-called 'great Australian adjective' – has failed the most basic of all adjective tests. Finally, a third test for adjectivehood is whether or not the word can itself be modified. All good adjectives can be strengthened by intensifying words such as *very*. So, as predicted, *shut that very rusty gate* is fine. We can't say the same for *shut that very bloody gate.* Here, once again, you get the literal sense of *bloody,* something entirely different from *shut that bloody gate. Bloody* fails again.

There are other adjective tests we could throw at the great Australian *bloody* and it would fail these too. So, what is the word class to which it belongs? Being an intensifying expression suggests that it is some sort of adverb. Saying something is *bloody good* is much like to saying that it's *very good.* But in fact, *bloody* and *very* are poles apart. For instance, while you can say *that's no bloody good* or *that's too bloody bad*, you can't say *that's no very good* or *that's too very bad.* Someone might ask you *how good was it?* You might reply *very!* However, you wouldn't reply *bloody!* The linguistic behaviour of these words is quite different.

The semantics is odd too. For one thing, words like *bloody* don't describe what you think they are describing. When you say *shut that bloody gate*, you're expressing disapproval, but it's not necessarily directed at the gate. It is the state of affairs that is lamentable (although of course the gate may have something to do with it), and the disapproval is smeared over the entire sentence. Emotionally laden taboo words, like *bloody*, are neither adjectives nor good adverbs – none of the traditional word classes accommodates them nicely.

9.10 Sentence types

As we have seen, languages vary regarding the extent to which they rely on word order to communicate grammatical relations. English has quite fixed word order, whereas other languages have much more flexibility in this regard. Languages with freer word order tend to vary the way material is presented to communicate information about a part of the clause that is in focus. For fixed-word order languages, variations in word order typically signal specific grammatical meanings. For example, consider the English patterns shown in Table 9.6. In the second column, 'everything else' in brackets indicates other elements that can potentially occur in clauses (such as objects).

Table 9.6 Word order, sentence types and typical meanings in English

Word order	**Sentence type and typical meaning**	**Example**
SV (everything else)	Declarative (makes a statement)	*My hovercraft is full of eels*
V (everything else)	Imperative (issues a command)	*Fill my hovercraft with eels!*
VS (everything else)	Interrogative (asks yes/no questions)	*Is my hovercraft full of eels?*
Wh-pronoun VS (everything else)	Interrogative (asks information questions)	*Where are the eels?*

These few simple examples show how a combination of different word orders, along with some specific word forms, can be used to express a variety of different messages. For example, English speakers form interrogatives by inverting the usual declarative word order (and including special pronouns in the case of *wh*-interrogatives). Changing the word order in this way is actually quite a rare strategy and largely confined to European languages – for example, German, like English, places the verb in initial position:

(9.125a)

[illegible]s.’

(9.125b)

[illegible]?’

Many languages show more than one possibility simultaneously. For instance, Yiddish, a close relative of German, combines two grammatical strategies: a sentence-initial question marker *ci* and a word order change from the declarative (and, like English and German, a change in intonation):

(9.126a)

[illegible] (falling intonation)

(9.126b)

[illegible] (rising intonation)

Other strategies used to communicate these different intentions in different languages include sentence-final particles and different intonation patterns accompanying the speech. For example, in Mali clauses are always structured in the same way and it is often only the intonation pattern that allows us to distinguish a statement from a question. Only in the case of information questions is there an additional indication of the intended meaning with the use of interrogative words in the clause. The differences are summarised in Table 9.7.

Table 9.7 Sentence types and intonation patterns in Mali

Sentence type	**Intonation**
Declarative	Gradually falling pitch
Informational question (use of interrogative words)	Gradually falling pitch
Yes/no question	Sharp rise on final syllable
Imperative	Slight rise on final syllable

Japanese uses the sentence-final particle *ka* to distinguish between a question and a statement, while the particle *yo* is used to signal emphasis, such as you might find in a command.

(9.127)

Nihonjin	*desu*	*ka*
Japanese.person	be	question

‘Are you Japanese?’

(9.128)

Okoru	*na*	*yo*
get.angry	don’t	command

‘Don’t get angry at me!’

In Table 9.6, we listed the most usual function for each of the structural sentence types. However, we emphasise with trumpet blast and fanfare that this is just the typical correspondence between form and function. Interrogatives can be used for many purposes other than posing questions – *Can I have a latte?* is a *yes–no* interrogative, but does not require an answer *yes* or *no*, and *Why don’t you shut up?* is definitely not an information-seeking question! Questions can also be asked without using an interrogative structure – the declarative sentence *It’s garbage night* could be a subtle way of asking someone to take out the rubbish bins. As we’ve seen on other occasions, we always need to distinguish form from function. In short, labels like declarative, imperative, interrogative and so on = grammar (syntactic category), and labels like directive, statement, exclamation, question and so on = meaning (semantic category). There is no one-to-one correspondence between grammatical form (sentence type) and its meaning. As we will see in Chapter 14, language has to be more flexible than that – and there is always so much that goes beyond the literal meaning.

The universality of these sentence types: Implications

Think how often in a day you are asked a direct question, either face to face or perhaps in some sort of survey. Interrogatives were once thought to be universal, but it is now clear that this is not the case. Linguist Bill McGregor (1990), for example, has described how the Aboriginal language Gooniyandi (spoken in the Fitzroy Crossing area of Western Australia) has no grammatically distinct class of interrogative. This doesn't mean speakers aren't able to ask questions, but they go about this in a very different way. Requests for information don't differ grammatically from statements, so speakers provide information but leave something unspecified in the expectation that the hearer might respond to and fill in the blank. Statements are also made with the desire that they will trigger some sort of agreement or disagreement on the part of the hearer.

Malngaya-nhingi
Malngaya-from somewhere
'From Malngaya he went somewhere'

Diana Eades has also discussed various indirect ways in which speakers of Aboriginal English seek information – they have a similar strategy of making a statement in the hope that a hearer in the know will (immediately or at a later time) confirm or comment on the information.

The educational implications of this are significant, since in schools the main method of extracting information out of children is by asking direct questions. Institutions such as the law and the media also consider direct questions to be the most effective way of eliciting information. But, what if a person's first language has nothing that corresponds grammatically to an interrogative? These are important differences between languages, and they have serious consequences in all sorts of areas of communication, as we discuss in Chapters 13 and 14.

9.11 Some final thoughts

Grammar (syntax and also morphology) is typically what students find the most challenging about starting linguistics. Unless you have been involved in learning another language, there will be lots of concepts and terms that are new to you. One of us (who shall remain nameless) nearly left linguistics after the first assignment (Japanese phrase structure rules!) because it seemed so challenging. Adding to the challenge is that, for a couple of generations now, grammar has not been a feature of the school curriculum. This was motivated by the idea that 'natural language learning' provides children with a strong linguistic context and they learn language by osmosis. This was the message all over the English-speaking world. No attention was paid to grammar or to the forms of language; as Dick Hudson writes, the school's contribution was purely 'to provide a rich linguistic environment'. Unfortunately, this has meant that children have been growing up for decades without a clue about how their language works (which can be handy knowledge to have, as we impressed on you in Chapter 2).

So, this chapter may have thrown a lot of unfamiliar concepts and terminology at you. Along with brand new terms, we have introduced different definitions for terms you probably thought you already understood (such as *sentence, word* and *subject*). Outside linguistics, there is an extensive non-technical vocabulary used by people when talking about grammar, but the terminology is often too imprecise to be of any real use for the systematic study of language and how it works, creating more trouble than it solves. By getting used to the terminology that is commonly shared among linguists, you will be taking a giant leap towards understanding how language works.

Further reading

There are many different books dealing with sentence structure, and they can take very different theoretical approaches. Accessible texts compatible with the approach we have adopted here include Brown and Miller (2002), and Tallerman (2015). Radford (2009) is a useful resource for developing skills in analysing constituency. Further reading on English grammar can be found in Radford (2004), Börjars and Burridge (2019), and Huddleston and Pullum (2005). For details of how question phrasing can influence the testimony of an eyewitness, see Loftus (1975, 1979). For a brief account of some of the different theories of syntax, see Chapter 4 in Finch (2003). Hudson has written extensively on the place of linguistics in the curriculum;

his papers on language and education can be viewed on his website (http://dickhudson.com/papers). The Japanese phrase structure rules are based on Fujinami (1996), but note that we have adapted them for our own purposes here. The Japanese sentence examples in the exercises section have been adapted from Nemoto (1999). The Māori examples are from Bauer (1993) and the Turung sentences from Morey (2010). Eagleson (1994) provides an account of the forensic analysis of personal written texts.

Exercises and discussion points

1 Grammatical relations in Japanese

Consider the following examples from Japanese. Each of these sentences is arranged in a different order, although they have essentially the same meaning in comparison with each other. They are best translated as 'Taro bought a book at that store'. What allows the meaning to remain stable even though the order of the words is changed? Complete the missing glosses based on your understanding of the meaning of the sentence.

[illegible]
[illegible]
[illegible]
[illegible]
[illegible]
[illegible]
[illegible]
[illegible]
[illegible]
[illegible]
[illegible]
[illegible]

---

2 Transitivity in Turung

What is the transitivity type of each of the following Turung sentences?

[illegible]
[illegible]
'They should sit at the head side'
[illegible]

[illegible]

‘The tiger ate the deer.’

[illegible]

[illegible]

‘Their house was near the river.’

(d) [illegible]

[illegible]

‘I am going to sleep.’

[illegible]mun²

[illegible]

‘I gave food [lit: rice] to the monk.’

3 Transitivity patterns in Mali

Although we have said that word order in Mali is basically SVO, there is some variation in the system. What word order patterns can you identify in the following sentences from Mali?

(a) [illegible]mes-ta

[illegible]man-human.plural

‘The men go hunting.’

(b) [illegible]ēsmēs

[illegible]food

‘They would carry food.’

(c) [illegible]ule

[illegible]

‘ … and he stayed there.’

4 Semantic roles and constituent order in Auslan

Using concepts from this chapter, including semantic roles, grammatical relations and ordering, explain how grammatical relations are organised in the following Auslan examples. Note that these examples do not provide enough information to describe the full Auslan system; they have been selected to indicate the range of options available.

(a) [illegible]

‘The boy eats cake.’

(b) [illegible]noun-3

‘She is buying a car.’

(c) [illegible] point]

‘The dog bites the cat.’

(d) table [left hand] boy [right hand] table-surface [left hand]human-jump-off [right hand]

‘The boy jumped off the table.’
(e) baby
‘The baby is crying.’
(f) dance pronoun-3
‘She is dancing.’

---

5 Complex sentences in Gurindji

Each of the following examples contains a complex sentence in Gurindji. For each example, identify the subordinate clause in Gurindji and state how it relates to the main clause.

[illegible]ayina
[illegible].on.them
‘I saw the old women when I was talking.’
[illegible]inyi
[illegible]
‘After the meeting finishes, I will come back.’
[illegible]ku
[illegible]-topic
‘That old woman went inside to talk.’
[illegible]ma-rna
[illegible]randmother
[illegible]
[illegible]me
‘I miss my maternal grandmother who gives me lollies.’
[illegible]-nga
[illegible]might
[illegible]-lu-ma-nyjurra-nga
crocodile-agent-topic-you.mob-might
will.bite‘If you mob go to the river, a crocodile might bite you.’

Research project

Analysing Yoda-speak

We noted in section 9.4 that Yoda-speak often involves moving phrases to the start of the clause (fronting). We also noted that this is a common strategy used by English speakers to emphasise particular constituents in the clause. Based on the corpus of examples provided here, prepare a description of Yoda-speak syntax, then answer the questions that follow.

(a) Friends you have there.

(b) Found someone you have.

(c) Then a Jedi will you be.

(d) A long time have I watched.

(e) Consume you it will.

(f) Sick have I become, old and weak.

(g) Unexpected this is, and unfortunate.

(h) Much anger in him like his father.

(i) When 900 years old you reach, look as good, you will not, hmm?

(j) Your father he is.

(k) Luke when gone am I, the last of the Jedi will you be.

(l) Tired I am, rest I must.

(m) You think Yoda stops teaching, just because his student does not want to hear? Yoda a teacher is.

(n) Yoda teaches like drunkards drink.

(o) Humility endless is.

(p) From the dark path, no returning there is. Forever, the direction of your life it dominates.

(q) So certain, are you?

(r) For my ally is the Force, and a powerful ally it is.

1 Describe the function and phrase type of the fronted constituents.

2 How does the fronting in Yoda-speak compare with fronting in Standard English? Is it more or less constrained than English? Is it consistent? If there are exceptions, are they rule governed? To answer this question, you will have to identify the conditions that trigger fronting in both Yoda-speak and English and identify where there are differences.

3 Using the Yoda-speak generator, we Yodified the various example sentences used throughout this chapter. You can find them in the online materials. Based on this extended corpus, what other elements of Yoda-speak, not relating to fronting, can you identify?

For more activities, please go to www.cambridge.edu.au/academic/linguistics.

Part 4

‘Language is a social fact’

Variation and change

10

Variation and identity

◆

> Language is a city to the building of which every human being brought a stone; yet he is no more to be credited with the grand result than the acaleph which adds a cell to the coral reef which is the basis of the continent.
>
> Ralph Waldo Emerson (1875), 'Quotation and Originality', in *Letters and Social Aims*

10.0 Introduction

People often think about language as fixed. If you play Scrabble and you want to know how to spell a word, you can look it up in the dictionary and there it is – the same as it was 20 years ago when the dictionary was first published. At the same time, people are also highly aware of the ways that speakers differ from each other and the fact that language can change. This causes problems for the 'language is fixed in time' model. For people who are committed to this view, change and differences among speakers become a signal of something that is not right – a lack of respect for and awareness of proper language. In this chapter, we will be looking a little at the issues to do with language standards, but mostly we want to explore all the different ways in which speakers align themselves with particular habits of using language. In some parts of the world, they can achieve this by switching between languages, but our focus here is on variation within a single language.

Never lose sight of the fact that language has two distinct functions. Its primary, or orthodox, function is of course to facilitate communication. The other is to promote identity, or in-group solidarity – to preserve a group's uniqueness or to establish its presence. The importance of this second function should never be under-estimated. Language is a potent symbol of a vast array of different affiliations. In addition to regional attachments, these include ethnicity, gender, race, sexual orientation, religion, generation, abilities, appearance and profession. Aspects of accent, vocabulary and grammar help to define the gang, whatever that gang might be.

10.1 ‘The chaos of a living speech’

To give you an idea of just how much linguistic variety there is out there, consider the following examples (spoken and written) from a language you all know well.

(10.1) *He'll should can come in the morn, will he no?*
Modern English (conversation, Hawick in Scotland)

(10.2) *The snottygobble* [= parasitic creeper] *is into this one, but ‘e’s not goin’ to kill ‘im.*
Modern English (conversation, rural Tasmania)

(10.3) *Each Christmas, at this time, my beloved father broadcast a message to his people in all parts of the world. Today I am doing this to you, who are now my people.*
Modern English (Queen Elizabeth’s first Christmas message, 1952)

(10.4) *We have got the world divided into boxes. each box is a science, and inside it we have locked. From the vast maternal quarry. Nature, And so, in little heaps assembled by chance or perhaps.*
Modern English (description on the packaging of a new microphone)

(10.5) *Where an offeree who has accepted a take-over offer that is subject to a prescribed condition receives a copy of a notice under sub-section (10) in relation to a variation of offers under the relevant take-over scheme, being a variation the effect of which is to postpone for a period exceeding one month the time when the offeror’s obligations under the take-over scheme are to be satisfied, the offeree may …*
Modern English (section 27(12) of the Companies (Acquisition of Shares) (Victoria) Code)

(10.6) *I lodges long on muster and mother Stewart, and I seed that ere lady a getting hup stares. I wur 'pon the bed, a'having a lie down, as his nat'ral nuff seeing has how master Stewart vos avay: cos, don't ye know, ven the cats avay, the mices does play*.
Early Modern English (mid-nineteenth-century eyewitness account from Sydney Police Office records)

(10.7) *Sod Egges: Seeth your Egs almost hard, then peele them, and cut them in quarters, then take a little Butter and put it in a platter upon your egges*.
Early Modern English (seventeenth-century recipe for boiled, or 'sodden', eggs)

(10.8) *Bipo tru, tripela liklik pik ol i stap long bush. Ol i no gat haus na ol i laikim wokim haus*. [Long ago (lit. 'before true'), three little pigs lived in the bush. They had no house and they wanted to build house(s).]
English creole (extract from a Tok Pisin story 'Tripela Liklik Pik' ['Three Little Pigs'], PNG)

(10.9) *So just now he asked me to rush back home first and clear all the stuff in my house lah, what … what … ever police thing that I got bring back home, the notes and so on lah. So he asked me to clear all this sort of thing lah …*
Modern Singapore English (transcript of a police-taped conversation)

(10.10) *<mark> didnt your motherboard come with any papers*

<Darkman-X> iz that that book that says A7V333 on it?

<[RaW]> yes

<Darkman-X> lol the one that i'm using to prop up my comp table?

<[RaW]> probably

<Darkman-X> whoops

<Darkman-X> :-)

<Darkman-X> j/k

<JoHn> lol

<[RaW]> yur supposed to use your school books for that dummy

Modern English (chatroom exchange)

All these different examples fall under the umbrella of ‘English’. Yet they are clearly very different – even those roughly labelled ‘Modern English’ (although in Chapter 12 we discuss the more controversial case of the English-based creole Tok Pisin). A shared history unites them (all have links of some sort with the group of continental Germanic dialects that ended up in the British Isles sometime in the fifth century AD), yet you can see some exotic-looking vocabulary here, as well as some very un-English-looking structure – for some passages, we even had to supply full translations. Linguistic labels like ‘English’ are misleading; they suggest fixed monolithic entities and obscure what, in reality, can be an astonishing range of variation.

The kind of variation we see represented here falls along two dimensions. The first is variation across time. As example 10.7 shows, time influences language. Even within one’s own lifetime, there can be changes in grammar, words and pronunciation, and if the timespan is long enough, the changes can be so spectacular that the language is no longer comprehensible. We won’t have much to say about this kind of variation here; it is covered in Chapter 11.

The second dimension is variation across space, and this includes both geographical and social distances. English is a language that now finds itself in most regions of the globe and this has changed the concept of the English **speech community** beyond recognition – no longer are we simply looking at a bunch of people bound together because they happen to live close by and speak the same language. Recent years have seen a burgeoning of diversity in the form of new contact varieties like Tok Pisin (example 10.8), New Englishes like Singapore English (example 10.9) and curious hybrids such as those found in global advertising-speak or product labels (example 10.4, where the unusual features are probably the result of machine translation). Even some native English dialects, especially the non-mainstream or ‘traditional dialects’, can have strikingly different structures; example 10.1 offers a taste of Scottish English and example 10.2 a snippet of Australian English from rural Tasmania.

Any socially significant group of people will differ in their linguistic behaviour – you would not expect the Queen of England (example 10.3) and a farmer in Tasmania (example 10.2) to speak in the same way. In addition, we can distinguish varieties on the basis of subject matter and domain; that is, the activity in which language plays a part. We therefore find different language used in the law (example 10.5), personal ads, government, linguistics, religion, sport and so on. We also have to consider whether the language is written or spoken, and the level of formality; a chatroom exchange (example 10.10) is very different from many other types of writing. Clearly, people talk differently in different contexts, too – there are differences in an informal chat, an interview, a lecture, a royal message and so on. Within all this variation there is even personal variation (the linguistic details of individual speakers, including social and geographical features, but also personality, physique and physical condition). No two people use language in exactly the same way, each of us is marked out as an individual by our unique linguistic behaviour.

Just imagine for a moment what any language would be like without this kind of variation. Indeed, from time to time people have tried to artificially create such linguistic systems – languages perfectly uniform and free of irregularities. But, even if such a constructed language were somehow to succeed and become the first language of speakers, it would eventually be 'struck down' by variability and indeterminacy. The simple truth is that a totally homogenous language would be dysfunctional. Think of the vast range of behaviours and situations that any human language has to cover – languages are culturally transmitted, and they will be used differently by speakers from different regions, from different social classes, of different ages, of different occupations, of different gender identification, of different sexual orientation and so on. Any label such as 'New Zealand English', 'Russian' or 'Pennsylvania German' is simply a convenient cover term for what are really clumps or clusters of mutually intelligible speech varieties.

10.2 Languages, dialects, accents – and standards

> A shprakh iz a dialekt mit an armey un flot. ['A language is a dialect with an army and navy'].
>
> Max Weinreich, Yiddish Scientific Institute Conference, 5 January 1945

When we start to discuss linguistic variation, we very quickly hit problems of terminology. Because terms such as language, **dialect** and **accent** are used in everyday English, they come with a lot of baggage, and these extra associations can easily prejudice or mislead. For many people, a dialect may bring to mind something rather quaint and rustic, while for others the term is quite pejorative. Dialects are varieties that don't come up to scratch – they are haphazard, vulgar or provincial aberrations of a 'real' (that is, standard) language. All sorts of things enter into people's judgements regarding what makes something a dialect and what makes it a language, and we need to shed this baggage if these are to be useful labels. Let's start with the technical sense of ‘dialect’ and ‘accent’ and how they differ from ‘language’.

Most people are aware of pronunciation differences – curiously, though, they tend to think of an accent as being something that other people have. In fact, everyone who speaks a language speaks it with some sort of accent. For instance, some speakers in North America pronounce their *r*s in words such as *car* and *first*, whereas most speakers in Australia or New Zealand do not (these varieties are non-rhotic). Both ‘R-fulness’ and ‘R-lessness’ are features of accent.

People are generally fascinated by vocabulary differences, too. Whether their favourite sausage is described as *Belgium sausage, Byron sausage, devon, Empire sausage, fritz, German sausage, luncheon sausage, polony, Strasburg* or *Windsor sausage* is undoubtedly significant for speakers – as an important part of their identity – but, like accent, these differences are not enough to establish separate dialects unless they are accompanied by grammatical differences of some kind.

So, what we need is something like the distinctive verbal system of Scots English illustrated in example 10.1, the so-called double (or even triple) modal construction. Standard English has a strict requirement that only one modal verb (such as *will* or *can*) appears in a sentence. In Scots English, and also some southern varieties of United States English, sentences like the following are possible:

(10.11) He might could do it.
(10.12) He should can do it.
(10.13) He’ll might could do it.

If you are struggling to understand what these might mean, you only need to unpack the modal verbs. *He might could do it* suggests ‘he probably could do it’; *He should can do it* means something like ‘He ought to be able to do it’. If you unload the meanings of all three modals in the last example, it becomes ‘It’s likely that he’ll be able to do it’. Subtle nuances of meaning can be captured very neatly by these compound modal constructions – and with time you’ll might could get used to it!

Dialects then are all the regionally, socially and also temporally defined varieties that fall within a larger named grouping like ‘the English language’, ‘the Russian language’, ‘the Japanese language’ and so on. While pronunciation differences alone don’t make a variety a distinct dialect, if speakers have a different accent, they will likely also show differences in their grammar. Although we have made a sharp demarcation here between accent and dialect, differences in accent are usually accompanied by distinct lexical and grammatical features.

Standardised language

Variety might be a fact of linguistic life, but there are homogenising forces at work seeking to stamp it out. In Chapter 15, we look more closely at what happens to languages when they are **standardised**, but really it isn't too different from normalising processes in other domains. You might compare the story of what happened to bananas when the European Union brought in regulations for their production. While crooked bananas weren't actually banned (this is a 'euromyth'), Regulation (EC) 2257/94 did specify that the fruit must be 'free from malformation or abnormal curvature'. Standardisation (international and national) promotes uniformity and establishes norms and requirements, whether it be for bananas, screw-thread sizes or languages.

In the case of a language, standardisation legitimises a single fixed and approved variety. So, dictionaries and grammar books record, regulate, tidy up and iron out variation. With the ground rules on usage laid down, there is no longer any official room for options. Speakers cannot vacillate between forms such as *I done it* and *I did it*, or *lie* and *lay* – one choice only carries the stamp of approval. All this can be done by deliberate and organised intervention or, as in the case of English, in a more haphazard fashion.

The English standard was artificially created over many years, not by any English language academy (because there never was one), but by a network of different groups, including writers of guides and manuals, dictionary makers, editors, teachers and newspaper columnists. Over the years, their collective endeavours have amassed an arsenal of prescriptive texts (dictionaries, grammars, style guides and so on) that have elevated and buttressed a single variety. That lucky variety happened to be London English, but if a city other than London (for example, Liverpool or Manchester) had had the same non-linguistic (economic, political and social) advantages, the socially prestigious dialect of that region would have been subject to the same elevation and spread. This is typically how standard 'languages' arise. It's all about context.

Something approaching the uniformity demanded by a standard is only achievable in the context of writing, especially the formal written language. The writing process (and, particularly in formal writing, the conscious self-censorship that accompanies it) insulates a language to some extent from the flux and variation that is the reality of any natural language system. There is also a sense in which the clout of the written word and the 'one-size-fits-all' ideology of the standard tend to blind speakers of standardised languages to messy linguistic reality. The neat lists, elegant definitions and finely spun paradigms you find in prescriptive texts look the other way and ignore 'the boundless chaos of a living speech' (as Samuel Johnson put it in 1755). Speech communities can be extremely complex, yet variability and mutability – qualities intrinsic to any linguistic system – don't sit happily within the classifications of a standard variety.

Miscommunication across dialects can put the wrong person in prison

The opportunities for misunderstandings based on dialect differences are considerable. Imagine the potential confusion when an Aboriginal child says to their teacher, *You gotta gimme paper? Gotta* in some varieties of Aboriginal English functions as a marker of future time (much like *gonna*), but a teacher unaware of this might well interpret the sentence as a directive, and a rather impolite one at that. Serious miscommunication has arisen in legal situations because the Aboriginal people involved have spoken a non-standard form of English and/or a creole lexically related to English – in other words, distinct linguistic varieties that are often not properly recognised or understood outside their speech communities. One case on which sociolinguist Diana Eades worked involved a man whose confession to a murder was fabricated by the police. Kelvin Condren maintained his innocence, claiming that he was 'verballed' in his police record of interview (so they were someone else's words not his). Eades could show that certain grammatical structures in his 'confession' were simply not part of his Aboriginal English. However, the judges ruled the linguistic evidence inadmissible and Condren was jailed for life. They had failed to grasp that this Queensland man could speak a variety of English that was related to Standard Australian English but nonetheless had very different grammatical rules. It took seven years before Condren's conviction was eventually quashed.

Another aspect to any form of standardisation is prestige. When the European Union brought in regulations for chocolate, chocolate products in the United Kingdom and Scandinavia were judged non-standard, and 'purists' even campaigned that these should be relabelled 'vegelate' or 'chocolate-flavoured' (to reflect the fact that they didn't contain the required proportion of cocoa butter). Whether applied to chocolate or languages, the label 'standard' entails not only 'uniform practice' but also a kind of 'best practice' – a measure of excellence (or 'benchmark') against which other varieties are then judged. However, prestige is a tricky notion. While varieties like the standard are associated with power, education and wealth (highly valued by many people), 'bad language' can have a prestige all of its own – a hidden prestige. For example, non-standard usage has 'street cred' for certain groups and, as attitude studies show, so-called low or powerless speech styles also win out when it comes to more human qualities such as integrity, social attractiveness and friendliness. In Chapter 15, we'll be looking more at this tug of love between the **overt prestige** of the high-status groups and the **covert prestige** of the local, non-prestige groups.

When do dialects become languages?

Earlier, we described language labels such as 'Australian English' or 'Russian' as convenient umbrella terms for clusters of mutually intelligible dialects. But, the whole concept of mutual intelligibility is a murky one. The small North Frisian dialect area in western Germany is a good illustration. Early descriptions divided the community into a number of individual village varieties that, by reputation, were mutually incomprehensible. It took the pro-Frisian language movement to unify this diversity. Suddenly, Frisian speakers from these different villages could understand one another, so psychological factors such as motivation and solidarity can encourage comprehension.

We encounter problems because notions such as language and dialect are not exclusively linguistic terms. Geographical, political and socio-historical concerns are also relevant – all are interrelated and none is immune to change. For instance, language differences are an important part of what distinguishes one nation from another and, over time, national frontiers can become linguistic frontiers. This means that some varieties will be thought of as separate languages simply because they come to be spoken in different countries. We think of Norwegian speakers and Swedish speakers as using separate languages, yet they can actually understand each other.

Another complication is that varieties can merge to produce a kind of cline or continuum of variation. 'Aboriginal English' involves such a cline, ranging from varieties at one end of the continuum that are no different from English elsewhere in Australia (except perhaps in accent) to those at the other end that are not mutually comprehensible. The Gurindji speech community beautifully illustrates the kind of the complexity typical of Aboriginal communities in rural Australia – traditional Aboriginal languages (for example, Gurindji) continue to be spoken by some people alongside other languages (such as Kriol, Aboriginal English and Standard Australian English), as well as new mixed languages (for example, Gurindji Kriol) that have arisen though the interaction of all these languages. These new languages range along a spectrum: at one end the mixed languages (with forms and structures of a number of languages), at the other an English variety fairly typical of rural Australia, and in between varieties of the English-based creole. Speakers move along this continuum, altering their speech to suit the situation and audience.

Features of Aboriginal English

One way of talking about the variation within Aboriginal English is to compare heavy and light varieties. Of course, there's no hard line between these varieties: people vary the way they talk depending on a range of factors, including who is present, their relationships and the purpose of the interaction.

In a paper summarising key features of Aboriginal English, Andrew Butcher uses the content in Table 10.1 to show how much closer the light version of Aboriginal English is to Standard Australian English than to the heavy version in terms of phonology. Focus here is on the sibilant consonants (the hissing and hushing sounds) in Standard Australian English such as /s/, /ʃ/ and /tʃ/ (though the examples also show how vowels vary between the varieties). In the heavy variety, you'll see that all the sibilants are pronounced [t̥ᶝ] (this is a type of voiceless (post)alveolar stop).

Table 10.1 Aboriginal and Standard Australian English

	Standard Australian English	**Light Aboriginal English**	**Heavy Aboriginal English**
'same'	[seɪm]	[sæɪm]	[t̥ᶝæːm]
'shame'	[ʃeɪm]	[ʃæɪm]	[t̥ᶝæːm]
'fish'	[fɪʃ]	[fɪʃ]	[pɪt̥ᶝ]
'face'	[feɪs]	[fæɪs]/[fæɪʃ]	[pæt̥ᶝ]
'cheeky'	[ˈtʃiki]	[ˈtʃiki]	[ˈt̥ᶝiki]
'catch'	[kʰætʃ]	[kætʃ]	[kæt̥ᶝ]
'picture'	[ˈpʰɪktʃə]	[ˈpɪktʃə]	[ˈpɪt̥ᶝə]

[illegible]

There is quite a lot of difference between the standard form and the heavy variety, and when we add grammatical and pragmatic differences into the mix, the challenges for intercultural communication are obviously significant. Grammatical features, for example, show significant differences in word order (*'E buy muticar, that blue-one Toyota* = 'He bought a blue Toyota') and in ways of presenting information (e.g. 'there is/are' constructions, as in *'e got plenty mussels* = ' … there are lots of mussels').

There are geographical clines too, sometimes called **dialect chains**. Here speakers in communities located nearby can understand one another (perhaps they live in neighbouring towns), but speakers located at the geographical extremes of these chains cannot. Dialect chains stretch across Europe. Near the German and Dutch border, for instance, German speakers using their own 'home-grown' varieties can communicate with neighbouring Dutch speakers better than they can with German speakers in the south of Germany, who speak southern varieties of German. So, where does German end and where does Dutch begin? There's so much variation and indeterminacy here that any attempt at dividing up such a continuum on the basis of purely linguistic criteria (such as mutual intelligibility) is problematic.

Despite all these complicating factors, we will adopt the usual linguistic rule of thumb: dialects are mutually comprehensible, whereas languages are not. (We will ignore for the moment that time messes with mutual intelligibility – Modern English speakers can't understand speakers from the medieval period, or beyond, but we still call it all English!) Just remember that individual labels such as 'German', 'Mali' and 'English' give the impression of easily identifiable and neatly compartmentalised entities, but such tidy classifications are nowhere near reality, and we will spend the rest of this chapter showing just how untidy this reality is.

10.3 User-related differences: Regional variation

> Whenever an Englishman would cry 'All right!' an American cries 'GO ahead!' which is somewhat expressive of the national character of the two nations.

Charles Dickens (1842), *American Notes*

A number of the differences in examples 10.1–10.10 were the fallout of both geography and society, and they involved variation associated with speakers (that is, user-related variation). At any given point in time, language will differ both between countries and within the same country; any socially significant group of people will also differ in linguistic behaviour.

Let’s start with linguistic variation associated with geographical location, since it is perhaps more obvious to people than some other types of variation. It is very apparent that a person in Manchester in the United Kingdom will have different linguistic patterns from someone living in, say, New York in the United States, Wellington in New Zealand, Kitchener in Canada, Johannesburg in South Africa or Brisbane in Australia. What we have here are different **regional dialects**. These can include national varieties, such as American English, Canadian English, British English, Australian English, New Zealand English and Irish English, as well as distinct regional dialects within these national varieties; examples of the latter include Boston English in the United States, Birmingham English in the United Kingdom and Waterloo County English in Canada. While such dialects can include both standard and non-standard speech forms, difference will always be more obvious between non-standard varieties.

Speakers of distinctive regional dialects will always have a distinctive regional accent as well, but as we described earlier, the opposite is not necessarily true. It is possible for speakers to have a regional accent, but their grammar will be similar to speakers from a much wider area. Standard English, for example, is used and understood by people all over the world and its grammar provides little or no information about regional background. This is largely true with respect to vocabulary too, although there can be some lexical items that are regional – there is nothing strikingly Antipodean about our use of grammar in this book, but the occasional vocabulary item might give the game away. However, if you heard either of the authors speaking, you’d have a pretty good idea where we come from.

Both Australian and New Zealand English are varieties that are geographically very uniform. This unity is the result of the original dialect mixing and the transience of speakers between regions. The original transported English was polydialectal, used by settlers from many parts of Great Britain and Ireland, but following migration, the speakers of these different dialects rubbed shoulders with each other regularly and the dialect differences were levelled out. Think of a giant linguistic melting pot.

Dialect melting pots down under

When dialects from all over the British Isles came together in Australia and New Zealand, a kind of fusion of linguistic features took place. In Australia, the original mix was largely made up of varieties from south-east England, Ireland and Scotland (in order of strength of input), with a particularly hefty dollop of London English. In a short period of time, all the different varieties spoken by convicts and free settlers underwent a general levelling process to produce a new compromise dialect, with the features of transported south-eastern British English figuring prominently. In New Zealand, links with Australia were well established before different varieties of British English arrived during the 1800s, so the 'colonial' children had established their accents before the overlay of regional varieties from the different parts of the British Isles. New Zealand would have experienced an intermingling of varieties and the subsequent blending predicted by the melting pot effect.

Many early observers of the speech of Australians and New Zealanders remarked on the uniformity and 'purity' of the language (and by purity they didn't mean the absence of 'foul language'!). Samuel McBurney writing for the *Argus* (Melbourne, Victoria) on 24 April 1886 described how he came across a class of boys and girls 'of English, Irish, Scotch, German and Tasmanian parentage to three generations' – yet, 'all spoke very much in the same way'. McBurney's accounts, strengthened by reports from all around Australia, tell of the levelling of regional features that was happening around this time – in his words, 'a new dialect is growing up'.

Case study: Variation in Australian and New Zealand English

Lexical differences tend to attract a lot of attention. Around Australia, a medium-sized glass of beer (approximately 285 millilitres) goes by many different names – in Melbourne a *pot*, in Sydney a *middy*, in Adelaide a *schooner* and in Alice Springs a *ten*. *Swimming costume* is used Australia-wide, alongside regionalisms such as *bathers* in the southern states, *swimmers* and *cossies* in New South Wales, and *togs* in Queensland and the south-eastern mainland. There are also expressions that appear confined to very specific locations: *hook turn*, 'a right-hand turn made from the left side of the road when the green light becomes red' is a peculiarly Melburnian driving manoeuvre. (You can see just how finely tuned this variation is by exploring the *Macquarie Dictionary Word Map*, an interactive resource that plots the growing lexical regionalisms for some 27 regional divisions.) As you'd expect, New Zealand also shows regional differences in vocabulary – for example, the northern-central-southern region is split between the use of *tiggy, tag*, or *tig* for the name of the children's chasing game. Research suggests that lexical variation of this kind has existed for a number of decades.

What about phonological differences? When it comes to accent in Australian and New Zealand English, popular claims that people can identify someone's place of origin purely on the basis of how that person speaks are exaggerated. Accent differences are still not particularly striking. It is more likely that they are a matter of statistical tendency, with a certain pronunciation occurring more in one place than another. For example, speakers from Hobart and Melbourne are more likely to say *graph* with [æ], while Sydney and Brisbane speakers are more likely to pronounce the word with [ɐ], with Adelaide and Perth speakers showing this tendency even more strongly. However, linguists have described this variation as complex. For one thing, the vowels don't occur uniformly across all words: throughout Australia there is a high usage of [æ] for *dance* and an extremely low incidence of this sound in *giraffe* and *mask*.

As yet, research has identified little in the way of grammatical variation that is regionally inspired. Probably the closest thing to an actual regional dialect is Southland English, a form of New Zealand English spoken in the deep south of the South Island. This was a pocket of settlement where Scottish English dominated, and consequently it has retained some Scottish-inspired features of syntax (such as *Did you not?* rather than *Didn't you?*, and *The garden needs watered* rather than the more usual *The garden needs watering*), pronunciation (notably the so-called Southland burr, a semi-rhotic feature) and vocabulary (*wee* 'small'). Because there are these grammatical differences, alongside pronunciation and vocabulary, we can describe this as a regional dialect.

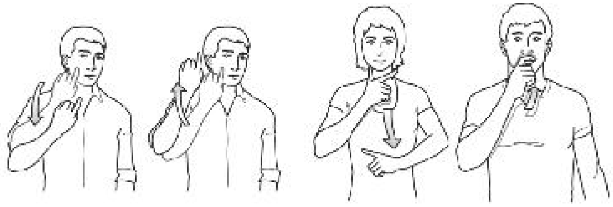
We expect that regional variation will become more obvious in both countries. As we describe in Chapter 11, all it requires are three magic ingredients: time, physical/social distance and the processes of linguistic change. English hasn't been in Australia or New Zealand for very long, certainly not long in terms of language change. Standardising influences aside, the distances between population centres (especially in Australia) are considerable, and regional chauvinism – as evident in the sort of strong rivalry between places such as Sydney, Melbourne and Perth – is a major incentive for people to start highlighting their distinctiveness linguistically. This is very evident in the way the Englishes in New Zealand and Australia are now diverging. The fact that there is no single prestige regional variety in either country also means that, if groups want to be identified regionally, varieties are free to go their separate ways.

For example, Andy Pawley has identified interesting grammatical features for what he labels Tasmanian Vernacular English. One area of interest is the appearance of gender marking on inanimate nouns (e.g. items of food and drink are always feminine: *I put* ***’er*** [= the bottle of beer] *down that bloody quick that I blew the top off* ***’er***; and masculine pronouns are used in reference to inanimate objects like trees – see example 10.2 above). Different ethnic mixes are also adding new dimensions to both Australian and New Zealand Englishes, in particular Māori English and Aboriginal English, varieties that grew out of the original contact situation between English and Indigenous languages. However, it is hard to unravel regional and social factors here. Varieties like Tasmanian Vernacular English typically coincide with particular socio-economic classes in speech communities, and ethnicity is a crucial part of social identity and something that people want to demonstrate through their use of language. So, it is tricky to talk about regionally defined variation here without appealing to aspects of class and ethnicity (see section 10.4 on social variation).

Case study: Regional variation in Auslan

In their book on Auslan, Johnson and Shembri make the point that signed language is often assumed to be a form of universal language. This is incorrect. Just as spoken languages can have very different phonological inventories, signed languages can be made up of different building blocks for creating signs. All around the world, there are many signed languages showing distinctive vocabularies (involving factors to do with handshape, location and movement of signs) and grammars (morphological, syntactic and discourse factors).

Auslan signers produce SISTER by tapping the handshape twice on the nose, and while the related languages British Sign Language (BSL) and New Zealand Sign Language (NZSL) share this sign, in American Sign Language (ASL) and Taiwan Sign Language (TSL) the sign is very different. As you can see in Figure 10.1, TSL has two signs to distinguish ELDER-SISTER and YOUNGER-SISTER. You can also see TSL has a handshape where only the ring finger is extended; this handshape is not available in Auslan.



Younger sister (TSL) Elder sister (TSL) Sister (ASL) Sister (Auslan)

Figure 10.1 The sign for SISTER in three different sign languages

The order of signs can also vary across languages. While the preferred order in Auslan has the actor preceding the verb and the undergoer following it (for example, SISTER SEE DOG), Japanese Sign Language has the verb at the end, but the same order for actor and undergoer (so SISTER DOG SEE); other signed languages, like ASL, have more flexible sign ordering. Negation strategies also vary. Auslan uses a headshake and Greek Sign Language uses a backward head tilt, while others draw on a different inventory of non-manual signals.

Even those sign languages that are related (such as Auslan, BSL and NZSL) have often diverged in their development, and within these languages there is regional variation. Currently, Auslan has two main regional varieties: a northern dialect encompassing New South Wales and Queensland, and a southern dialect including all the other states. These varieties show lexical differences that have come about via formation changes involving signs for numbers, colours, temporal terms (such as YESTERDAY, AFTERNOON, LAST-WEEK) and also question signs (such as WHO). Such variation also interacts with additional factors like age – for example, differences are less obvious in younger signers.

Clearly, different languages may not change in the same way when they find themselves in different locations, and we expect regional variation will continue for Auslan, just as it will for Australian and New Zealand English. Johnson and Shembri report that signers in Perth, Brisbane and Adelaide are less likely to lower signs such as NAME, THINK and CLEVER than signers in Sydney and Melbourne; for instance, Perth signers produce NAME at the forehead, while Sydney signers produce it at the jaw. Such differences could increase over time and eventually lead to more significant regional variation.

10.4 User-related differences: Social variation

> It is very vulgar to talk like a dentist when one isn't a dentist. It produces a false impression.

Oscar Wilde (1895), *The Importance of Being Earnest*

The kinds of regional variation that we have been describing are often mixed up with other variation, especially of a social nature. In the case of the [æ]–[ɐ] variation, age and gender don't seem to be involved, but the [ɐ] variant is definitely perceived as belonging to a higher social class. It is also more formal, especially for words such as *dance* and *plant* with [nasal + fricative / stop] (so it is probably not a coincidence that this is the vowel that occurs in *la-di-da*!). Such matters of style have to do with the way we alter our language to suit the occasion, and we address this use-related variation in section 10.5. Here we are focusing on variation related to speakers, in particular their background and their social group; in other words, we are looking at **social dialects** (or **sociolects**). In both Australia and New Zealand, there are differences that classify us into specific groups. They include those marked by level of educational achievement and type of occupation. These differences are reflected in the way we speak English – in the way we use sounds, our choice of vocabulary and our grammar.

Often, social factors are closely entwined and it is impossible (and probably not even useful) to tease them apart. In the Auslan community, for instance, there are some significant age-related differences. While some are the consequences of technological change (for example, the three signs for TELEPHONE reflect changes in the device itself), other differences interact with religion and education; they also relate to regional factors because of the way education of the Deaf was carried out in each state of Australia. In the past, states had large, centrally run schools for the Deaf, but Catholic children were educated in different schools and were taught different signs based on Irish Sign Language (ISL). With the phasing out of ISL in the 1950s, these linguistic differences have become associated with older signers in the community. Variation between the way men and women sign is also mixed up with age. In elderly groups, men are more prone to use fingerspelling with their Auslan (in other words, they use more hand configurations to represent the letters of the English writing system); because males used to have greater educational opportunities and work experiences than women, this gave them a better practical knowledge of English than females in the same age group. Additional influences come from a signer's family background, and the degree and nature of their exposure to English. As Johnston and Schembri describe, native signers (those who have learnt Auslan as a first language from Deaf parents) tend to use a type of sign language that has a more complex grammar compared with those who have learnt the language as a late-acquired first language in early adulthood, perhaps after limited or unsuccessful exposure to English. While these differences might begin as the fallout of the acquisition processes, they can easily become **social markers**, or features that identify signers as members of a particular social group.

To understand this kind of variation, we need to study the social lives of speakers, in particular the relationships between individuals and their membership in **social networks**. In their day-to-day lives, people interact with others in a number of different roles. For example, they might have relatives who are also neighbours and some of these individuals might even share a workplace. This sort of interaction involves what are called dense networks; the greater the number of links between the same individuals, the denser the networks are said to be. As we see in Chapter 11, such communities are great norm enforcers – speakers with few external contacts will identify with and orient themselves towards those with whom they interact most extensively. These networks correlate with conservative speech patterns and a lack of innovation. On the other hand, where there is little social cohesion between members of the speech community, social networks are said to be loose; such networks are more typical of larger, often urban communities with many external contacts, and they are more open to innovation.

Case study: Age-related variation

We can't do justice to all the **social variables** with which speakers align, so we will focus on one in particular here: age. Of course, all sorts of things happen to our speech as a result of the ageing process. These are changes that most of us (well, those of us around a certain age) don't really want to contemplate – decreased rib mobility (and accompanying decline in respiratory function), calcification of laryngeal cartilage, changes to muscle tissue, bone loss in the upper and lower jaw (resulting in reconfiguration of the vocal tract itself). Let's not concern ourselves with speech variation that comes about through these sorts of anatomical changes. What we are looking at is age as a social division, and therefore something that people wish to demonstrate through their use of language.

Age-related variation involves all sorts of **linguistic variables** we could consider . A number of them are highlighted in the following extract from a letter of complaint written by 70-something Vivienne Player, and published on 4 September 2012. We will focus on two.

Fellow humans with whole new language

> Like, I'm on the train, right? It's, like, 9 o'clock on Saturday morning and it's f–ing freezing? Like nine degrees or something? And these two, like, chicks get on and their skirts are right up to their, like, bums? And their white legs are, like, naked? And covered in, like, chickenbumps?
>
> At the next station, this real cool dude with a backpack swings on and they both, like, take their ear phones out? One girl is like 'Oh my God, it's Cam!' and the other's, like 'Cool!' Cam sits down and he's, like, all over them and they, like, seem to think he's really, like, hot? They both, like, hang on his every word, even though he's only got, like, about 10 of them? And the top five are: 'f–', 'shit', 'cool', 'massive' and, like, 'like'?
>
> When they get off at, like, Richmond, çOh, my God! Cam's backpack gives me a, like, massive thwhack in the face as he passes my seat? And he, like, doesn't even notice? (I'm, like, a woman in my 70s?) Hey, guys, don't forget we older generations, like, exist? Even if we are another species and speak a different language.
>
> Vivienne Player, Beaumaris

High rising terminal

The question marks in Vivienne Player's letter suggest a prosodic feature that is stereotypically associated with youthspeak. It goes by various names, including the high rising tone (HRT) or terminal and uptalk. Although the use of HRTs also occurs in the United States, Canada and Britain, it has been stereotyped (and often stigmatised) as a distinctive pattern of Australian and New Zealand English since the early 1960s. Given the greater frequency of HRTs reported in studies conducted in New Zealand (which is where the feature was first spotted), it is possible that the use of HRTs has been imported from New Zealand into Australia. Then again, it might always have been in Australian English, but just have been slower to take off. Unfortunately, there is a ten-year gap between the first systematic data collection in the two countries and methodologies of the various studies differ substantially, so it is impossible to know for sure.

The following are some actual examples, both from radio talkback; the arrow (⇑) indicates where the rising tones have occurred.

(10.14) My parents come from Greece ⇑ but I was born in

Melbourne. ⇑

(10.15) My name's Maree ⇑, I live in Richmond ⇑ and I'm a hairdresser. ⇑

According to studies of the HRT in both Australian and New Zealand English, the social evaluation criteria attached to this phenomenon are as follows:

- associated with females and young people
- expresses uncertainty, lack of confidence and deference on the part of the speaker, yet they are perceived as more friendly and attentive to the listener's needs
- disapproval from others for excessive use.

A number of studies of the HRT in both Australian and New Zealand English have since tried to sift out the stereotypes. Overwhelmingly, they conclude that the use of this distinctive intonation does not express uncertainty, but rather seeks verification of the speaker's understanding; in this sense, the speaker is seeking to accommodate the addressee without being markedly deferential or uncertain. It is similar in use to hedging expressions (such as *you know, right?*) and it promotes speaker–hearer interaction. Studies also show that the most frequent users of HRTs in both countries have been young women (although males do use it, and older speakers too!), and in New Zealand it is more prevalent in Māori English. As Edgar Schneider reports, this distinctive intonation appears to be spreading globally, especially among younger speakers; it is now being reported in British, American, and Caribbean Englishes, and occasionally also in Africa and Asia.

Discourse markers

Another feature in Vivienne Player's letter is the hedging expression *like*. Its literal meaning, 'similar, resembling something', has given rise to various colloquial uses in discourse and her letter (though a parody of youthspeak) illustrates some of these. First, *like* can be a focus marker – 'I'm, like, a woman in my 70s'. Second, it can function as a hedge that reduces or weakens the force of an utterance, as in 'like nine degrees or something'. A third *like* is something that has been dubbed the 'quotative *like*', a use that first appeared in the late twentieth century and, similar to the HRT, appears to be on the increase. It introduces reported speech, noises, facial expressions and also thoughts or attitudes – ‘the other’s, like, “Cool”’.

Celeste Rodriguez Louro has studied the quotative *like* in Australian English. Her study of quotation devices in contemporary Perth English (spontaneous narratives of personal experience recorded in 2011) makes an interesting comparison with the findings of an earlier study on Melbourne English in the late 1990s, and suggests an increase in the use of *be like*, particularly among the youngest speakers. The results are in line with trends that have been noted in the literature on English quotation elsewhere in the English-speaking world. In New Zealand English, the quotative *like* and discourse marker *like* are frequent, especially among young speakers.

The linguistic landscape on Palmerston Island

The work that Rachel Hendery has been carrying out on Palmerston English shows the extent to which speakers will align themselves with particular ways of using language. This variety of English is used by a tiny isolated community in the Cook Islands (its total land area is only around 1 square mile, or 2.6 square kilometres); in August 2009, its population numbered only 54. Given its isolation, it is perhaps not surprising that over the past 140 years it has developed a unique linguistic and cultural identity influenced by England, the Cook Islands and, more recently, New Zealand. Yet, what does seem extraordinary is that even this tiny atoll with its tiny speech community shows linguistic diversity. As Hendery describes, the island is divided three ways, and each third belongs to one of three families: Akakaingara, Matavia or Tepou (named after the three Cook Island wives of William Marsters, the man who settled the place in the mid-1800s). Cutting across the divide between the three families is another distinction, that between the ‘beachfellas’ (who live in the cleared sandy main settlement area at the north end of the island) and the ‘bush people’ (who have built houses among the palm trees further south). While the actual physical distance between these groups is short (a walk of a few minutes between even the most far-flung houses), the social distance is considerable, and linguistic variation is the fallout – phonological, morphological and even syntactic features distinguish the language of the beachfellas and the bush people.

Case study: EthnicITY-related variation

Ethnicity is another aspect of linguistic variation that reflects the social dimension of language usage. **Ethnolects** are associated with certain ethnic or cultural sub-groups, and they may be marked by phonological, syntactic and lexical variation (usually influenced by speakers' first languages (or L1s) or those of their families). Māori English and Aboriginal English are important ethnolects that grew out of the original contact situation between English and Indigenous languages, and in New Zealand there is a new ethnolect on the block – Pasifika English, which we describe below. These three ethnolects are important markers of ethnic group membership, and are often employed as in-group codes in addition to general Australian and New Zealand English.

Varieties of Aboriginal English differ from general Australian English at all linguistic levels. As described above, accent shows a continuum from a 'heavy' accent (close to the sound system of traditional Aboriginal languages) to a 'light' accent (close to the sound system of Australian English). Lexical differences can be striking: some words are either borrowed directly from Aboriginal languages (*goona* for 'shit' may have this history) or have changed dramatically (*gubba* 'white man', probably from *governor*); familiar-looking English words can have quite different meanings (for example, the future marker *got to/gotta, sorry business* 'ceremony associated with death' and *deadly* 'fantastic'); while some early English words are also maintained (*gammon* 'joking, pretending', eighteenth-century underworld slang). The grammar of Aboriginal English has plenty of features that are found in other non-standard dialects, but it also has many of the very different grammatical features that appear in contact varieties like Kriol (e.g. transitive marking: *-em /-im* on verbs: *We seeim buffalo got big horn* 'We saw a buffalo with big horns').

New Zealand now has well over 200 000 people whose families immigrated from the South Pacific islands, largely from the Cook Islands, Tonga, Samoa and Niue. As these communities shift from their Polynesian languages to English, there is the potential for speakers to start signalling cultural and ethnic identity through a distinctive form of English. Pasifika English shares features with Māori English, and a number of them are spreading beyond the Pasifika communities. Though the exact relationship between Pasifika English and Māori English has not yet been established, a number of features are known to be shared – for example, /ð/ pronounced [d] in words like *there*; /θ/ pronounced [f] in words like *birth*; unaspirated stops in words like *pub*; /r/ pronounced in words like *girl*; and syllable timing (the 'rat-a-tat-a-tat' rhythm described in Chapter 7 versus the 'tee-tum-tee-tum' quality of standard English). Some of these features are clearly transfers from Polynesian languages and a number are acquiring social meaning. Moreover, as linguists Bell and Gibson (2008) describe, they are being picked up by younger speakers, not just Pasifika and Māori people, but also Pākehā (New Zealanders of European descent). Important here is the significance of Māori English as a solidarity marker and the practice of 'language crossing' (sometimes referred to as 'bro talk'). Speakers are no longer confined to those who are ethnically Māori – Pākehā growing up or identifying with Māori groups may also speak the variety, and as Bell and Gibson conclude, the expansion of these ethnolectal features is very important to the overall development of New Zealand English.

10.5 Use-related differences: Situational variation

We now move to the type of variation that is associated with occasions of use (that is, situational variation). Language varies constantly in response to many different situational factors, including the relationship between speakers and their audience, the setting and the subject matter. An important type of contextually defined variety involves something called a **register**. Language is tailored to certain situations, occupations or subjects – for example, you will find that there are registers associated with groups such as doctors, lawyers (as in example 10.5), airline pilots, advertisers, bank managers and musicians, particularly hip-hop artists. These varieties are not tied to particular speech communities in the way that (social or regional) dialects are, yet they can be just as distinct from each other and, like dialects, are differentiated at all levels of language: phonetic, lexical, grammatical, discourse, semantic and pragmatic. Accounts of registers frequently characterise them solely in terms of their distinctive vocabularies; perhaps the focus on lexicon is understandable – unfamiliar words and phrases are always much more obvious to people than grammatical or phonological novelties, and of course these new expressions also name the things that are especially significant for speakers. However, as the following examples illustrate, registers involve much more than simply lexical differences.

(10.16) *SOLID Scorpio mum wants friend, soul mate, likes motorcycles, mountains, meaningful man, camping, exploring festivals, blues, kids, seeks any age, looks, background for future in grass roots lifestyle!* [Personal ad, *Melbourne Times*]

(10.17) *A fast-medium right arm inswing bowler needs two or three slips, a deep third man, a gully, a deepish mid-off, a man at deep fine leg and another at wide mid-on.* [cricketese]

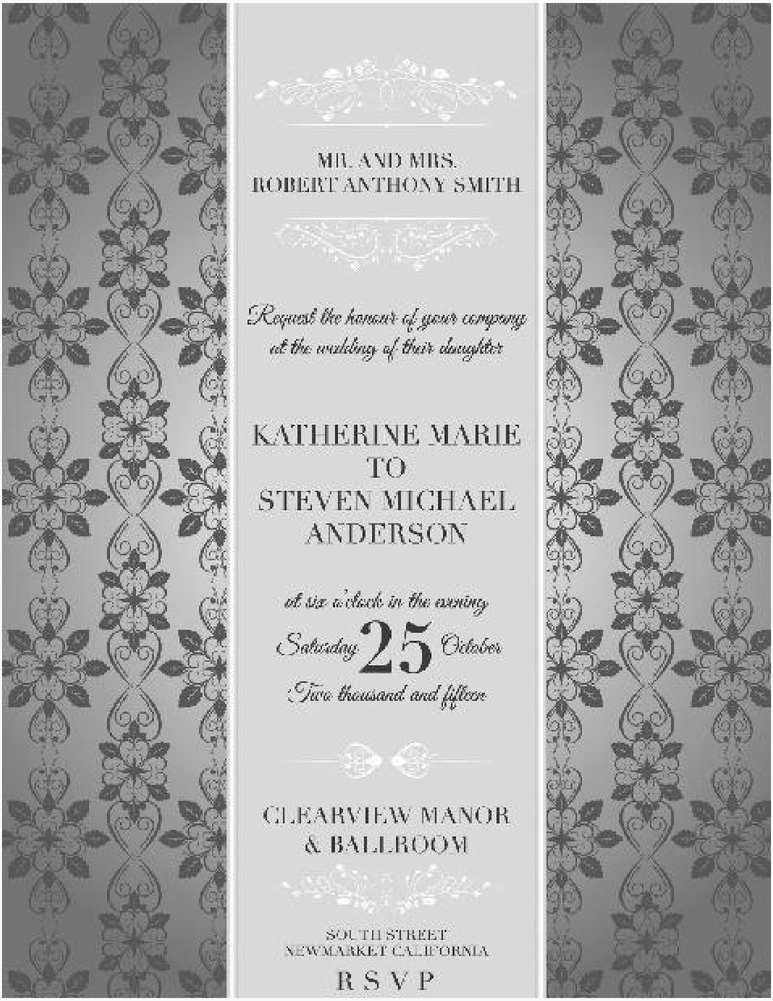
(10.18) *Cast on 63 sts: Knit 6 rows plain knitting. 7th row: K4, wl. fwd. K2 tog to the last 3 sts. K3* . [knitting instructions]

(10.19) *A term of a sale shall not be taken to exclude, restrict or modify the application of this Part unless the term does so expressly or is inconsistent with that provision.* [Legal English, or legalese]

(10.20) *Gud retro cust svC!!!* [tweet]

(10.21) *One dollar bid, now two, now two, will ya' give me two? Two dollar bid, now three, now three, will ya' give me three?* [auction chant]

(10.22) [formal invitation](Figure 10.2).



MR. AND MRS.
ROBERT ANTHONY SMITH

Request the honour of your company
at the wedding of their daughter

KATHERINE MARIE
TO
STEVEN MICHAEL
ANDERSON

at six o'clock in the evening
Saturday 25 October
Two thousand and fifteen

CLEARVIEW MANOR
& BALLROOM

SOUTH STREET
NEWMARKET CALIFORNIA
RSVP

Figure 10.2 Formal invitation

As these few examples indicate, we can characterise specialist language use by certain criteria, including the following.

- *Lexical markers:* distinctive vocabulary specialised for use in some particular domain – for example, the idioms and abbreviations of personal ads.
- *Phonological markers:* distinctive characteristics of speech – for example, the hushed tones of 'funeralese', and the speed and fluency of sporting commentaries and auction chants.
- *Presentational markers:* distinctive typographical conventions and layout – for example, the special format of an invitation.

- *Syntactic markers:* distinctive grammatical structures – for example, imperatives in instructional texts (*knit 6 rows*), agentless passives and impersonal constructions in academic discourse (the texts are marked with syntactic complexity), and full noun phrases rather than pronouns in legal documents (the repetition of *offeree and offeror* in example 10.5).

All these features work together to create distinct registers for different groups.

Register more or less coincides with *jargon*, but doesn't have the same negative associations. As a more technical term, a 'term of art', **register** is not generally used in ordinary language, so has escaped deterioration. Outsiders can be put off by language that is not readily comprehensible, so most people have two very different senses for **jargon**: (1) the language peculiar to a trade, profession or other group; and (2) unintelligible or meaningless talk or writing – gibberish. If you are inside the group, then specialist language is efficient, economical and even crucial, since it can capture distinctions not made in ordinary language; however, if you happen to be outside the group, the language can be unintelligible and alienating. The material in this book has introduced a host of special terms (such as *lexeme, schwa* and *polysynthetic*), and these have made it easier to talk in precise terms about language. You have also encountered ordinary language items with specialised meanings, such as *accent* and *dialect*. People without some background in linguistics may feel excluded and consider the terminology to be pretentious, but this applies to every field of learning, and indeed every walk of life – knitters, carpenters, wine connoisseurs, surfies, medicos and linguists all have their specialised language.

So, registers are varieties that we use when we are 'on the job', so to speak. Unlike varieties that are regionally or socially determined, they do not involve 'permanent' linguistic fixtures. Like the other 'lects' we have considered so far, however, they mark the boundaries of the in-group and the out-group by distinctive features of language. Solidarity is conveyed through these social passwords. It could be the Latinisms of medicos and lawyers, the slang of ecstasy users, the special colour vocabulary of interior decorators and so on.

Talking pink in the playground

A nine-year-old girl we know, who isn't really into pink and sparkly things and who takes a lot of cues about the world from her teenage brother, mentioned one day that she had taken on the role in her class of interpreting the boys for the girls and vice versa. Some of what she described is associated with the different concerns that groups of kids (mainly organised along gender lines) have in terms of sharing space in the playground and resolving conflicts associated with wanting to do different types of activities (soccer versus milling around and talking) in the shared space. Yet, the differences run deeper than that. The 'girl' groups, of which she does not consider herself to be a member, are conversant with and preoccupied by relationships and status. They manage conflict through (sometimes extended) negotiation. Navigating this terrain, our informant tells us, requires the ability to 'talk pink'. 'Pink language' sounds a lot like a register, with its extended resources in relation to topics such as social media (*WhatsApp*), grooming (*fishtail braids*) and peer relationships (*besties*). On the other hand, the 'boy' groups (which also do not attract her in terms of membership) seem to experience conflict as brief but explosive bouts of shouting and jostling that generally have no ongoing impact on relationships. Vocabulary choice and language practices in these groups include a lot of big talk (*Your mother* …) and mimicking the speech of hyper-masculine role models such as hip-hop artists. It is interesting to notice just how thoroughly gendered children are by the age of nine, and how these patterns are reflected in both the form and content of the language they use.

10.6 Style: What makes individuals and groups distinctive?

> Style is life! It is the very life-blood of thought!
>
> Gustave Flaubert, letter to Louise Colet, 7 September 1853

Style is an important aspect of language that is determined by use; or, in other words, by where language in a particular context falls on the spectrum of (in)**formality**. Consider the following exchange between Basil Fawlty and Mr Hutchinson, two characters from the comic sitcom *Fawlty Towers*.

Mr Hutchinson: There is a documentary on BBC2 this evening about 'Squawking Bird', the leader of the Blackfoot Indians in the late 1860s. Now, this starts at 8.45 and goes on for approximately three-quarters of an hour.

Basil Fawlty: I'm sorry, are you talking to me?

Mr Hutchinson: Indeed I am. Yes, now, is it possible for me to reserve the BBC2 channel for the duration of this televisual feast?

Basil Fawlty: Why don't you talk properly?

(*Fawlty Towers*, Series 1, 'The Hotel Inspectors')

Much of the humour of the Mr Hutchinson character in this episode derives from the fact that he speaks in an extremely elevated and florid fashion. He sounds like a book and it is all bizarrely out of place in his everyday interaction with Basil Fawlty. The language also clashes amusingly with his accent, which is not the posh BBC-accented English that normally accompanies this formality.

When creating language, speakers and writers select particular forms in response to the degree of (in)formality and familiarity appropriate to the context. We can identify an up–down continuum from formal through to informal varieties. Formal varieties, even those involving very different groups of people, tend to be more similar to each other than informal varieties; this is because language associated with formal settings usually involves the **standard language**.

There are different ways of dividing up this continuum of formality. One popular model was put forward by Martin Joos in the 1960s. He identified five styles of English:

- **frozen** – the most formal and polished style, usually reserved for important (perhaps large public) occasions and using formulaic language
- **formal** – a high style, used for important occasions but with less sophisticated language than the first style
- **consultative** – involves two-way participation (for example, doctor–patient, lawyer–client, teacher–student); not overly formal but words are chosen with care
- **casual** – relaxed conversational language, as used between friends, colleagues and some family members
- **intimate** – private language used for in-group interaction as between close family members, lovers and very good friends.

We illustrate these levels in Table 10.2 using simple address terms (ranging from greater to less formality). You will notice that the boundaries between these different styles are by no means clear-cut and for many people the use of titles and surnames is more than a little formal. In addition, both casual and intimate styles can be described as colloquial. They are both more usually spoken and they assume a fair amount of shared knowledge or common ground between the participants.

Table 10.2 Joos' five styles

Frozen (or oratorical) style	Madam Chair, Mr Chairman
Formal (or deliberative) style	Madam, Sir
Consultative style	Mrs Jones, Mr Jackson
Casual style	Esme, George, Jonesie, Jacko
Intimate style	darling, snookums, bae

There are no fixed boundaries between each of the different styles, and any one person's language will reflect a wide range between these extremes of frozen and intimate style. Style will vary according to: who we are and with whom we are communicating; whether we are speaking or writing; where we are and when the utterance takes place; what we are talking about; and how we feel about the whole situation. Change any one of these factors, and the style may well change accordingly.

To illustrate, consider the domain of the law and the register used by lawyers (often dubbed legalese). Like most varieties, this occurs in a range of styles. 'Frozen' style occurs in written documents like wills, but can also be spoken in the case of a witness's pledge to *tell the truth, the whole truth, and nothing but the truth*. Written 'formal' style is found in statutes and briefs, and can be spoken in the courtroom in the arguments of counsel and the examination of witnesses. 'Consultative' style is usually dialogue and used in lawyer–client interaction and lay witness testimonies. 'Casual' style removes the social barriers and can usually only occur during in-group interaction (for example, in the conversations between lawyers), while 'intimate' style would involve close friendships (or perhaps kinship), probably away from a work setting. These represent different varieties of the register of legal English. They all have in common some of the linguistic features of legalese, but they differ in their relative frequency and the special combination of these features. Style is always a matter of linguistic choice, even if many of the choices are made without much conscious reflection.

The Rillington Place murders

> It is good for the linguist to know that he can be useful.
>
> Jan Svartvik (1968), *The Evans Statements: The case for Forensic*

Throughout this book, we present a number of examples of forensic linguistics, the branch of linguistics that is neatly situated at the intersection of language and the law. One very early example of forensic linguistics at work is the so-called ‘Rillington Place murders’ – you might be familiar with the case from the 2016 TV mini-series.

Despite the fact that he had always claimed he was innocent, in 1950 Timothy John Evans was hanged for the murders of his wife and child. At the time, Evans had accused John Reginald Halliday Christie of murdering his family, even though the police had produced verbatim record where Evans had supposedly confessed to police that he had committed these murders. When a number of people campaigned on Evans’ behalf (including several members of parliament), Swedish linguist Jan Svartik was brought in to examine the evidence. In 1968, Svartvik showed that the confession used to convict Evans was most likely the product of police influence and differed sharply in style and structure from the remainder of Evans’ statements. Evans was illiterate, yet aspects of his ‘confession’ showed an uncharacteristically formal style, especially for a speaker of his background: ‘she [his wife] was incurring one debt after another’; ‘in a fit of temper’; ‘the 12.55 am train’ (Evans’ usual way of reporting time was ‘the five to one train’). Evans was posthumously pardoned. When several more bodies were discovered at 10 Rillington Place, Christie was found guilty of the crimes and then hanged.

Vocabulary will always show the most impressive range of styles, as Table 10.3 illustrates.

Table 10.3 Five levels of formality

[illegible]
[illegible]
[illegible]
[illegible]
[illegible]
[illegible]
[illegible]

When we put together a table like this, we are imagining some sort of default context by which we can suggest a stylistic value of an expression without reference to any context of use (this is something like the etiquette gauge that dictionary editors use when they brand expressions for offensiveness with labels such as ‘insulting’, ‘offensive’ and ‘taboo’). Even so, you may not agree with our classifications. Connotations differ (sometimes wildly) between speakers. Another complication is that the differences may not always be a matter of pure style; meanings can interfere (for instance, *elderly* has a feature of ‘seniority’ that *old* lacks). Recall the discussion on collocation in Chapter 6 – this is another complicating factor.

We should point out that certain **semantic fields** are distinguished by their relative formality. For instance, we had no trouble coming up with intimate or casual expressions for vomiting or drunkenness. In this up–down scale of (in)formality, these fields are ‘bottom heavy’. They draw overwhelmingly from the stock of native English words and are characterised by a flourishing of low-level slang expressions. There are literally hundreds of these expressions, such as *retch, heave, spew, upchuck, barf, yarf, urp, hurl, ralph, puke* (for ‘vomit’) and *smashed, sloshed, stewed, pickled, stoned, tight, potted, tanked, legless* and *loaded* (for ‘drunk’). By comparison, *regurgitate/be nauseated* and *inebriated/intoxicated* (based on medieval Latin) are among the very few elevated terms.

Other semantic fields are characteristically ‘top heavy’ and draw from higher style expressions (usually of French and classical origins). We had no trouble coming up with formal terms such as *indigent, impecunious, destitute* and *impoverished* for ‘poor’, and one of the reasons for this is that poverty is an area of social taboo. This makes it an inevitable target for **euphemism** and ensures a long chain of obscure vocabulary to avoid saying that dirty word *poor*. Recent times have seen a rise in such dainty circumlocutions as *economically marginalised, negatively privileged, economically non-affluent, culturally deprived* or even *differently advantaged* – all French or classically inspired.

Stylistic variation can involve phonological features such as stress (for instance, PER*gola* compared to *per*GO*la*), or how distinctly (or not) a speaker articulates the sounds. In casual everyday speech at the informal end of the scale, we find higher levels of fast-speech phenomena such as the weakening and reductive changes we saw in Chapter 8 (for example, a flapped /t/ between vowels in words such as *latter* or *better*). When we are chatting with people we know well, in a familiar and relaxed environment, there is a lot of shared ground, and we can make do with less phonetic and semantic detail. So, we take short cuts!

Grammar is also involved, of course. Sentences such as *It is I* and *The data are misleading* will be used in completely different contexts from *It's me* and *The data is misleading*. Imagine if someone said *It is I* on the phone to a good friend. They would be laughed at or accused of putting on airs! However, there are contexts – say, a formal piece of prose, where formality such as *It is I who am responsible* would sound quite reasonable. Many people avoid something like *The data are misleading*, except perhaps in a piece of serious academic writing where the word *data* is used as a plural with a singular *datum*. Formal style is also characteristically 'nouny' (in other words, characterised by Latinate nominalisations) – for instance, *in the event of default in the payment of* versus *if you don't pay what you owe*. The sentence *The decision has been made to continue the utilisation of pulverulent egg products by reason of a fund insufficiency* is unlikely to be part of casual conversation between two friends chatting about cooking, but rather an announcement made by the CEO of a large restaurant chain. This is because the passive construction *has been made* is part of formal register, as are expressions such as *utilisation, pulverulent* and *fund insufficiency*. A more appropriate utterance between two friends would be *We're broke so we've decided to use powdered eggs*.

Speech styles in Japanese

The ways speakers create relative formality or informality vary across speech communities. Japanese is famous for its complex system of honorifics, with attention to politeness and formality saturating every aspect of the language, especially vocabulary and grammar. Different methods of conjugation, specialised vocabulary and the addition of honorific elements combine to indicate the relative status of the speaker, the listener and also any other persons mentioned. For example, the past tense normal polite form of the verb *motsu* 'to hold/have' is *mochimashita*, the casual form is *motta*, the exalted form is *omochi ni narimashita* and the humble form is *omochi shimashita*. The past tense normal polite form of *suru* 'to do' is *shimashita*, the casual form is *shita*, the exalted form is *nasaimashita* and the humble form is *itashimashita*. In some cases, the verb is even replaced by an entirely different word: *taberu* 'to eat' becomes *meshiagaru* in the polite form and *itadaku* in the humble form. In the case of nouns and adjectives, there is either different vocabulary or an honorific prefix is selected (*o-* or *go-*); for example, *jōzu* 'skilled' becomes *ojōzu*.

At least four levels of politeness have been identified, with some curly intricacies added: colloquial language for very familiar friends or 'talking down to someone' (this includes extremely brusque, provocative or rude forms, which would be considered below 'colloquial'); casual language for family or in-group member use; normal polite language for use by strangers or by adults who are not part of the same in-group – it can be used when 'talking up'; and honorific (or formal) language to show respect or emphasise social distance. Honorific forms may be humble (*kenjōgo*) or exalted (*sonkeigo*) forms, depending on whether you are 'lowering' (humbling) yourself or 'raising' someone else's stature.

Slanguage and swearing

The best illustrations of variation and identity come from colloquial language – in particular, its slanguage. For example, Polari (aka Parlare, Pa(r)lary and Parl(y)aree) is a **slang** that was used by British homosexuals in the twentieth century; it probably derives from slang used by actors and other show people, and overlaps with slang still widely used in the London area (and perhaps elsewhere). Another slangy variety is Leetspeak (or l33t), the symbolic writing of 'netheads' like online video gamers, and for a short while it enjoyed more widespread popularity. Examples of Leet can still be found on the internet, although this form of slang is probably well and truly passé by now (some describing it as having 'jumped the shark'!).

People are often rather conflicted about slang, and over the years the descriptions have been many and varied. Here is a taste:

> Slang is a necessary aspect of our linguistic well-being.
>
> Walt Whitman, *Slang in America* (1885)

> Slang is to a people's language what an epidemic disease is to their bodily constitution.
>
> John F. Genung (1893), *Outlines of Rhetoric*

> Slang is vigorous and apt. Probably most of our vital words were once slang.
>
> John Galsworthy (1927), *Castles in Spain and Other Screeds*

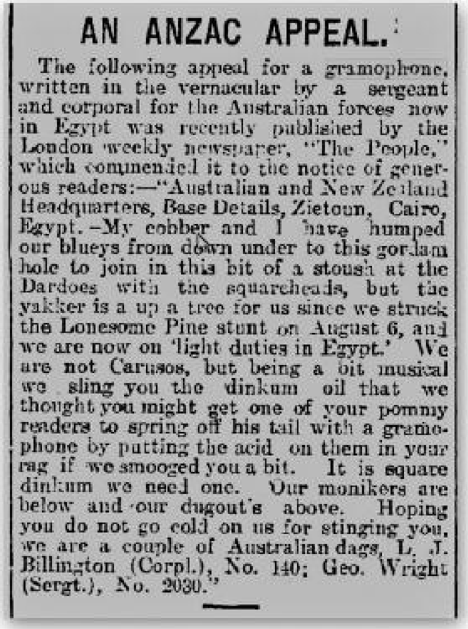
> [Slang is] dustbin language.
>
> Robert Burchfield (1985), *The English Language*

> Slang is language that rolls up its sleeves, spits on its hands and goes to work.

Carl Sandburg, *New York Times*, 13 February 1959

> Slang is a poor-man's poetry.
>
> John C. Moore (1962), *You English Words*

Many people connect slang with swearing, but this is wrong. Certainly, slang can be rude, but it is not connected to taboos in the same way that swearing is. Others wrongly identify slang with adolescent speech – as something most people eventually grow out of. Of course, teenagers use slang, but so do adults. Most importantly, though, they never use each other's slang! Finally, there are also those who view slang as a linguistic feature that is peculiar to modern times. This most certainly is not the case. Admittedly, it is tricky to make stylistic judgements on early language, but slang has always been with us – it just changes. The following news report in Figure 10.3 shows wonderful slang from 'a couple of Australian dags' from the 1910s.



AN ANZAC APPEAL.

The following appeal for a gramophone, written in the vernacular by a sergeant and corporal for the Australian forces now in Egypt was recently published by the London weekly newspaper, "The People," which commended it to the notice of generous readers:—"Australian and New Zealand Headquarters, Base Details, Zietoun, Cairo, Egypt.—My cobber and I have humped our blueys from down under to this gorlam hole to join in this bit of a stoush at the Dardoes with the squareheads, but the yakker is a up a tree for us since we struck the Lonesome Pine stunt on August 6, and we are now on 'light duties in Egypt.' We are not Carusos, but being a bit musical we sling you the dinkum oil that we thought you might get one of your pommy readers to spring off his tail with a gramophone by putting the acid on them in your rag if we smooged you a bit. It is square dinkum we need one. Our monikers are below and our dugout's above. Hoping you do not go cold on us for stinging you, we are a couple of Australian dags, L. J. Billington (Corpl.), No. 140; Geo. Wright (Sergt.), No. 2030."

Figure 10.3 'An Appeal for a Gramophone'

Slang mostly involves vocabulary, but pronunciation can be slangy too (*thing* rendered as *thang*), as can grammar (the invariant tag *innit*, as in 'weather's great, innit'). Other slangy bits of grammar are the stranded 'not' at the end of the sentence ('What a totally amazing, excellent discovery … NOT!!') and new uses of *because* ('She squished her feet regardless into pointy shoes because fashion'). We see here the kind of reduction typical of casual electronic communication, but in fact it is more interesting than this. *Because* appears before noun phrases (as in the previous example), as well as adjective phrases ('I've been missing out on sleep because addicted') and verb phrases ('I missed the ending because fell asleep').

Slang is used for a lot of reasons but its most important function is to act as a 'clique' or in-group recognition device, so it is a correlate of human groups with shared experiences, such as being children at a certain school or of a certain age, or being a member of a certain socially definable group, such as sex workers, drug addicts, jazz musicians, lawyers or professional criminals. Slang expressions will identify activities, events and objects that have become routine for those involved, and they have an important function in creating rapport in the work (or play) environment. These characteristics of slang overlap with defining characteristics of register, or jargon. Both slang and register work much like masonic mortar to stick members of a group together – at the same time erecting barriers between them and the outside. Try to identify examples of slang in your own speech and that of your friends. How would you feel about others using these terms?

Where slang is associated with activities in conflict within the more established sectors of society, the language has the added motivation of secrecy. Within the **anti-languages** of 'underworld' subcultures, we see slang expressions used to conceal the nefarious nature of whatever they designate, with the express purpose of keeping outsiders in the dark. The following is a handful of examples from the more than 2300 street terms that refer to illicit drug types or drug activity (taken from a glossary compiled by the White House Office of National Drug Control Policy): *get snotty* ('to use heroin'), *candy sticks* ('marijuana cigarettes laced with powdered cocaine'), *gym sticks* ('steroids used by athletes'), *lipton tea* ('poor quality drugs'), *klingons* ('crack addicts'). The high turnover rate of this vocabulary means that these terms are probably already out of date; the need to maintain secrecy ensures that as soon as a term's cover is blown, it has to be replaced.

It is quite clear that people take great delight in their slangy creations, and a striking feature of slang is its playfulness. In 2013, the word *selfie* (shorthand for a *self-portrait photograph*) was dubbed the *Oxford Dictionary*'s 'Word of the Year' (its frequency had increased by 17 000 per cent from the previous year). People then started to play with it, creating *helfie* (picture of one's hair), *belfie* (picture of one's posterior), *welfie* (workout selfie), *drelfie* (drunken selfie) and so on. As we noted in Chapter 4, 'the intensely pleasurable sensations generated by using, hearing or coining a new word or phrase' have even spawned a new word: *neologasm* is now listed in the *Urban Dictionary*.

Creeping colloquialisation

Written language and public speaking used to fall towards the ‘frozen’ and ‘formal’ end of the stylistic spectrum, but they are becoming progressively more casual and everyday in their nature. This textbook is itself a good example of this informalisation. The language we are using here is far more laid-back and much more personal than anything you will find in earlier textbooks (at least, up until the 1960s when the changes began). Take, for instance, the rather chatty way we constantly refer to you, the readers. If reference was ever made to reader(s) in the past, it was typically done in what now comes across as very stilted English, often using the third person – for example, ‘The astute young student [that’s you!] will doubtless have noticed the familiar style that has been adopted in this present book.’ Here’s an actual example from an English grammar book published in 1886 (but used into the 1960s):

> It is earnestly hoped that the slight sketches of the History of our Language and of its Literature may not only enable the young student to pass his examinations with success, but may also throw him into the attitude of mind of Oliver Twist, and induce him to ‘ask for more’.

Changes to terminology are also revealing: *lectures* are now more likely to be called *talks*. And when did you last hear labels like *oratory* (the art of public speaking), *rhetoric* (the art and study of persuasive writing and speaking), *elocution* (the art of public speaking where qualities of voice production, gesture and delivery are emphasised), and *recitation* (the act of reciting memorised materials in public)? These terms are simply no longer part of most people’s active vocabulary.

Swearing, like slang, is restricted to colloquial styles (which is not to say it never occurs in formal situations). In English, it includes religion-based profanity (for example, irreligious language), as well as a wealth of obscenities taken from the pool of ‘dirty words’. Only taboo terms can function as swear words. For instance, learned words for sexual organs and effluvia generally do not (**You faeces!*, **Urine off!*), and nor do certain mild obscenities and nursery terms (**You* *willie! *Wee-wee on you!*).

There are at least four (overlapping) functions for swearing.

- The **expletive** function. Expletives are a kind of exclamatory interjection that, like other interjections, have a highly expressive function (for example, *Wow! Ouch! Shit!*). They are uttered to let off steam in intense emotional situations.
- The **abusive** function. Swear words can add extra insult to injury. They include curses (*Bugger you!* or *Bugger off!*), personal slurs aimed at a second person (*You rotten bugger!*) or a slur made in relation to a third party or situation (*The bastard stole my pen* or *Writing grant applications is a bugger*).
- The **stylistic** function. Swear words are used to spice up the discourse, making utterances more noteworthy than if ordinary language had been used (Tourism Australia's invitation to the rest of the world: *So where the bloody hell are you?*).
- The **social** function. This is the use of swear words to display in-group solidarity (especially when directed against outsiders), as a part of verbal cuddling or simply friendly banter; it also covers expressions of mateship and endearment, such as *he's a good old bastard*.

Tō teka! Bullshit! Scheiße! Khuinya! Kuso!

A number of experiments have measured the emotional impact of words using techniques like electrodermal monitoring (a kind of polygraph). This records 'skin conductance responses' (an increase in sweat gland activity). Research supports overwhelmingly what every native speaker intuitively knows: 'dirty' words compared with other words evoke stronger skin conductance responses. Studies also confirm that swear words have more potency, more pungency and more wounding capacity in languages acquired early in life, and that the emotional force diminishes for languages learnt later in life. A study of English–Turkish bilinguals, for example, found that of all semantic categories, the greatest emotional reactivity was to taboo words in both languages, but the responses were much stronger in swear words in the speaker's first language than their second. These findings support what second-language speakers have often reported anecdotally – namely, that the expressive force is always much stronger in first-language swear words than second. Moreover, those who don’t like swearing find it easier to utter taboo words in their second language than their first (which for some can offer a freedom from first-language norms of behaviour). Studies also record that the auditory stimuli elicit greater emotional arousal than the visual stimuli in the first but not in the second language. In other words, the sound of taboo words is more disturbing than their appearance in print, once more confirming the subjective reports of language users generally: dirty words ‘sound awful’ and are much harder to say than to write.

10.7 Online identities

> I’m an internet linguist: I explore the language of the internet, for the people of the internet
>
> Gretchen McCulloch, https://gretchenmcculloch.com

In our all-digital-all-the-time age of Facebook, SnapChat, WhatsApp, chatrooms, web forums, online games, blogs, tweets and so on, we need to consider online identity before we finish this chapter. The internet has created a brand-new society where people converse with each other in virtual communities just as they do face to face in real communities. On social networking websites, members can express their individuality in many different ways. Apart from their profiles, people can add friends and personal information (birthdays, educational and career details, political views, religious beliefs, interests and so on), with links connecting these interests to pages that enable members to chat with like-minded people online. There are thousands of these interest groups.

The electronic revolution is not just a technological and social revolution, but a language revolution too. There are new expressions (*tweologisms,* such as *twitterholic, twaddiction, celebritweet* and *twitterati*), new conventions of spelling (alphanumeric texts like *l33t*, pronounced ‘leet’) and punctuation (emoticons and the Japanese *kaomoji*), and even new grammatical structures such as the ‘because X’ construction mentioned earlier (beautifully illustrated in the title of Gretchen McCulloch’s new book *Because Internet: Understanding the New Rules of Language*). These features herald the rise of new registers and the emergence of specialist varieties of electronic communication, alongside already existing varieties such as legalese and journalese.

The rapidly changing nature of this kind of language means that our examples will in all likelihood be ‘old hat’ by the time you are reading this textbook, and we recommend you seek out online pieces on evolving internet language (such as Gretchen McCulloch’s daily blog).

The informal and interactive nature of electronic exchanges (synchronous or asynchronous) allows for self-expression through creativity and linguistic playfulness, as people seek to mark themselves with in-group codes. The language can be very distinctive – much more speech-like, sharing many of the features not just of spoken language, but of actual conversation. There is slang (*graff* ‘to do graffiti’), swearing (*make me so fkn mad!!!!!*) and grammatical informality (*bestest everest friend*) – all the things normally frowned upon in formal writing. This kind of writing is full of the omissions, contractions (*dno* ‘don’t know’) and creative spellings (*matez* ‘mates’, *wateva* ‘whatever’) that we normally associate with the short-cuts and assimilations of ordinary rapid speech. The spontaneity and speed of such communication also mean that writers don’t go in for the same careful organisation and planning as they might with normal writing (*yer so im not deleting this coz i spent a while typing it*). Constructions are much looser, and they give us the expressive opportunities in word order that we typically associate with spoken language (*I get called a lot of oder names coz my name is kinda hard to pronounce so yeah I dont mind wat ever*), comment clauses and asides (*oh by the way just so you know*), interjections (*so wats happening with u huh*), filled pauses (*Umm about me*), and even repetition, false starts and digressions. All of these non-standard features help to build strong online identities.

In many ways, online language is reminiscent of the language we find in earlier times, when people simply wrote as they spoke – a time when there was no standard language and no prescriptive ethos telling writers what to do. In electronic communication, we see forms of regional and social variation, even personal variation, where people use punctuation, spellings and abbreviation to reflect their personalities. This is not surprising. We are looking at virtual speech communities where people communicate in writing. The only way members can express their identity – personal and shared – is to break free of the straitjacket of the standard.

Forensic linguistics: Perpetrating electronic identities

There are all sorts of implications of communicating with people without actually seeing them. In these virtual communities, people can ‘say’ whatever they want and whenever they want.

Away from face-to-face contact and within the safe confines of their internet personas, they are more likely to dispense with politeness paraphernalia and express views that may not be very ‘nice’. But, the growing field of forensics linguistics now means it is easier to track anonymous perpetrators simply by looking at their linguistic habits (vocabulary choices, morphological features and so on). People’s idiolects leave behind valuable clues.

So, you find that linguists are sometimes called in to provide authorship analyses in criminal investigations, involving cases of murder, rape, blackmail, extortion and the multiple identities of online paedophiles. In a discussion on the news website *The Conversation*, Tim Grant, the Director of the Centre for Forensic Linguistics at Aston University, describes how linguistic experts expose technological anonymity by gathering up the linguistic evidence that people leave behind as they write:

> Everything from the way someone uses capitalisation or personal pronouns, to the words someone typically omits or includes, to a breakdown of average word or sentence length, can help identify the writer of even a short text like a Tweet or text message.

In a 2009 murder trial, Grant contributed strong linguistic evidence to a criminal case against a man who had murdered his wife. He had tried to cover his tracks by sending text messages from her phone to colleagues and family members (including himself), hoping to give the impression she was still alive on the day he had killed her. Forensic analysis of the writing showed that the wife’s messages were not in keeping with her usual style but were more compatible with that of the husband.

10.8 The ‘chaos of a living speech’ is not so chaotic

Of course, people have always realised that different social classes and different situations triggered different linguistic behaviour, but it was considered messy variation and linguists didn’t want to be bothered with it. It was the American linguist William Labov who showed for the first time that these differences could be quantified. In his 1966 study of New York English, he started out by identifying a number of linguistic features that were used variably within the New York speech community – such as post-vocalic [r] in *beer* and *pork*, his most famous study. He then interviewed a random sample of the population, trying to elicit a range of different speaking styles – from informal conversation, through to a more formal interview style, to reading styles, the most formal of all. By doing this, he was able to show a clear interaction between speakers' social background and the formality of their speaking style. All the linguistic features he looked at varied significantly according to the social group of the speaker. For example, particular accents were especially associated with speakers higher up on the social scale, and all speakers adjusted their usage towards these more prestigious accents in their more formal styles. When people felt the situation becoming more formal, they pronounced the words in ways they perceived to be more 'correct'. In Labov's study, women appeared to be the most sensitive to the prestigious forms; on the whole, they showed greater shifts to the prestigious patterns in their formal styles.

Wherever you find similarities of social structure within a speech community, you can expect to find similar patterns of distribution for linguistic variables, and Labov's methods showed that earlier views of language variation had been incorrect – socially relevant variation is directly observable and can be studied in a systematic way. This new way of looking at language, variation and change sparked a whole new line of research into linguistic behaviour.

Researching language variation

> [S]ocial pressures are continually operating upon language, not from some remote point in the past, but as an immanent social force acting in the living present.
>
> William Labov (1963), 'Social Motivations of a Sound Change'

One of the tricky things about researching language variation is that it is one of those areas of life where what people *actually* do isn't always a particularly good fit with what they *say* they do. This is because quite a bit of our language behaviour bubbles along under the radar of our consciousness – although we can make active choices to speak in certain ways, we also make a lot of less-than-conscious choices about this as well. (That's quite aside from the more general problem (discussed in chapter 3) that people tend to report their ideal behaviour in preference to their actual behaviour – even when they have noticed what they do!) This is hardly surprising given the cognitive load in spoken communication. In addition to listening to the other person, and worrying about what we will say next, as we show in Chapter 16, we are also working hard on finding all the right morphology and arranging things in grammatical sequences before utterances pop out of our mouths and into the world.

As just described, William Labov pioneered research that explored the unconscious adjustments people make in their speech and built connections between language variation and the social context. By using observational methods and a quantitative approach, Labov was able to unpack the dynamics of language change (something we explore further in Chapter 11), showing how variation happens in real time in local contexts.

In another famous study, Labov uncovered the social motivations for a sound change taking place on the island of Martha's Vineyard (perhaps known to you because it was the site for the filming of the movie *Jaws*, and is now a popular holiday destination). He focused on unpacking the changing pronunciation of the diphthongs /ai/ and /au/; in particular, these diphthongs were observed to be centralising – the [a] at the start of the diphthong was moving towards [ə]. Labov identified three components of this sort of puzzle:

- the origin of linguistic variations
- the spread and propagation of linguistic changes
- the regularity of linguistic change.

Focusing especially on how changes spread, Labov emphasised that 'one cannot understand the development of a language change apart from the social life of the community in which it occurs'. Martha's Vineyard was a useful place to start on this puzzle because it was such a small and distinctive community. His research showed that this diphthong change was concentrated in particular communities on the island, both in terms of geography and social identity. In fact, the people who centralised the two diphthongs most markedly were members of the groups of islanders most resistant to the changing demographics of the island and least accepting of the tourist trade (around 40 000 visitors each summer versus the 6000 locals!). In negotiating islander identity, the centralisation of the diphthong can be seen as a means of asserting a Vineyard orientation.

Among the parameters that may contribute to sociolinguistic variation are age, social class, gender, style and register, mobility in space, mobility in class social network, particular professions or shared activities, ethnic or cultural group, and engagement with particular segments of the mass media. Here is a little more information about how to go about conducting variationist research.

Developing a project

Picking your data question

In a sociolinguistic variationist study, you need to know what sort of variation you are focusing on and why. What aspect of language are you concerned with? What are you hypothesising that you will find? How will you test your hypothesis? Keep in mind that any aspect of language may give you something to explore: sound change, morphological change, syntactic change, semantic change or the use of pragmatic features.

Picking your speech community

Sociolinguistic variation research seeks to understand and comment on the ways in which language is used in particular speech communities. Labov notes that variationist research works well in contexts where speech communities are clearly bounded, with 'well-defined limits, a common structural base, and a unified set of sociolinguistic norms'. In identifying a speech community to explore, think about what sets it apart from others and what is distinctive and shared within the group.

Collecting the data

Much of the early research on sociolinguistic variation focused on the observation of random participants in interaction. It is also possible to build more strategically focused corpora by seeking out participants with specific social characteristics. Sociolinguistic variationist data typically rely on the collection of extended interviews, in which participants are encouraged to tell stories about exciting, dangerous or memorable events. As discussed in Chapter 3, the choice of topic is strategic, encouraging people to focus on what they are narrating rather than the way they are speaking. The interviews are recorded, transcribed and annotated so that features of interest are visible to researchers. Other types of data that have been used in this research include observational data in places like shops, and corpora developed from the vast stores of written language available via the internet.

Analysing the data and interpreting the results

Not only does the variationist approach emphasise the descriptive orientation (outlined in Chapter 1) by observing people rather than asking them about what they (think they should) do, it also pulls the researcher out to a broader view by focusing on larger data sets in more abstract ways. Remember too that, as we discuss in Chapter 3, quantitative data on its own can answer questions about *how much* and *how often*, but not questions about *why* or *what does it all mean*. We recommend that you visit this chapter for details of research methodology textbooks that outline the ways in which quantitative evidence can be interpreted.

Further reading

For an overview of research into linguistic variation, see Tagliamonte (2012). On the formation of postcolonial dialects, see Trudgill (2004). The *Australian Journal of Linguistics* published a special issue devoted to English in Australia (2003, vol. 23, no. 2); Lawrie Zion's *The Sounds of Aus* is a 2007 documentary on the Australian accent, and for a more detailed account see Cox (2006), as well as Bradley (2004). For accounts of New Zealand English, see Bell and Kuiper (2001), Marsden (2007) and Hay, Maclagan and Gordon (2008); for Pasifika English, see Bell and Gibson (2008). On the sounds and grammar of Aboriginal English, see Malcolm (2004a, 2004b) and Butcher (2008). On Tasmanian Vernacular English, see Pawley (2008). Schneider (2004) looks at phonetic and phonological variation in English worldwide. Rodriguez Louro (2013) examines the use of *like* down under; Hendery (2012) describes the remarkable linguistic diversity on Palmerston Island; and Johnston and Schembri (2007) examine the regional and social variations in Auslan. For studies of the HRT, see Britain (1992) and Szakay (2008) for New Zealand English, and Horvath (1985), and Fletcher and Loakes (2006) for Australian English. In addition to the *Macquarie Dictionary Word Map*, we recommend the early surveys on language use in different parts of Australia carried out by Bryant (1997). *An Appeal for a Gramophone* was posted by Adrian Smith on the Facebook Page for the Gramophone and Phonograph Collector's Downunder (10 March 2018). Milroy and Milroy (1985) is an important work on language standardisation. See Labov (1963, 2007) for studies of the social motivation for change. You can find the whole of Svartvik's (1968) report online. Eades (2012, 2013) shows the serious implications of linguistic differences, and Grant (2013) describes how linguists track perpetrators by looking at the linguistic clues they leave when they chat online. Joos (1961) divides the continuum of formality into the five different stages we described here. Allan and Burridge (2010) discuss cusswords in Australian English, and in (2006) they describe the experiments that show the arousing and evocative nature of these words; Dewaele (2010) offers the multilingual perspective on swearing, and for slang see Adams

(2009). Gretchen McCulloch's piece on evolving internet language can be found online (https://gretchenmcculloch.com); see also her book, McCulloch (2019).

Exercises and discussion points

1 Variation in Auslan

Use an online dictionary to identify signs for the following items.

- **(a)** red
- **(b)** blue
- **(c)** yellow
- **(d)** yesterday
- **(e)** afternoon

Note how many of these items have regional variants and describe the variations.

2 Variation in Antipodean English

A striking feature of Australian and New Zealand Englishes has been their regional uniformity (especially in accent and grammar); however, this is now changing. Describe your own pronunciation of the following words (use phonetics); compare it with that of your classmates. Often regional variation is mixed up with social and stylistic variation. Is this the case here?

- **(a)** The vowel in *plant, dance, graph, grass, advance, example, castle*.
- **(b)** The vowel in *shell–shall, dole–doll, full–fool, pill–pull*.
- **(c)** The vowel in *trap, dress, kit*.
- **(d)** Pronunciations of verb forms such as *grown, shown, flown* (for example, [groʊən] versus [groʊn].
- **(e)** Whether the [j] is dropped in *nude, news, tutorial, enthusiasm, issue, assume*, presume.

Can you think of other examples that show variation in pronunciation?

3 English around the world

Have a search of the internet for examples of English used in countries where it is typically spoken as a second language (for example, China and Japan). Try something like ‘Japanese English images’ to find examples of signs.

(a) What types of examples did you find?

(b) What sorts of functions are these uses of English intended to address?

(c) What sorts of differences are there between your examples and more standard varieties of English?

(d) What may have motivated or caused these differences?

4 Style – levels of formality

As we saw, vocabulary will show the most impressive range of styles. Consider the following table and, as much as you can, find examples to illustrate the other styles. Make note of where you are unable to supply an example, and try to explain why this might be so. (Also make note where you disagree with our description of the words.)

[illegible]
cute
to drink
demented
affrighted
intelligent
toilet
conflict
copulate
peg out

5 Advertising

Collect around ten advertisements (you might want to confine yourself to one product – for example, food or cars). Take account of the features we’ve discussed in this chapter and describe both the attention-seeking devices and the tricky manoeuvres in the language of these advertisements.

Research project

A global standard for English?

In order to see how globally uniform Standard English is, the English linguist David Crystal describes in his *Encyclopedia of the English Language* how he collected samples from 40 English newspapers from countries all over the world. They were collected on the same day, 6 July 1993, and covered the same range of topics. He claims that on purely linguistic grounds, it was impossible to tell which paper belonged to which country. There was next to nothing in the way of an obviously regional use of grammar. The only occasional clue was in the form of a distinctive lexical item or spelling. His study suggested that, at that time, a worldwide Standard English was already in place – at least for writing. Replicate Crystal's study and discuss your findings.

For more activities, please go to www.cambridge.edu.au/academic/linguistics.

11

Structure across time

◆

> There can never be in language, just as there can never be in the ceaselessly blazing thoughts of men, a moment of true standstill.
>
> Wilhelm von Humboldt (1836), *On the Structural Variety of Human Language and Its Influence on the Intellectual Development of Mankind*

11.0 Introduction

Historical (or diachronic) linguistics is the branch of linguistics that studies the shifts and mutations of languages over time. To set the scene, consider the following lines from a fifteenth-century courtesy manual (a book that prepared adolescents for polite society). With a bit of phonological imagination, all the words can be unravelled, but their shapes, and the ways they are combined, are strange.

(11.1)

Ley not þyne Elbowe nor thy fyst
Lay not your Elbow nor your fist
Upon the table whyles þat thow etist.
upon the table while you eat
Bolk not as a Bene were in thy throte,
Belch not as (if) a bean were in your throat
[*The Lytylle Childrenes Lytil Boke or Edyllys be*, 1480]

Alas, we have no time machine to give us a glimpse of the accents of this period, but there are ways of reconstructing early sounds and sound systems. As a result, we know that the pronunciations of even familiar words were very different. The vowels in *bean* and *thy* were more like [ɛː] and [əiː]; the *wh* spelling represented the consonant cluster [hw] (which has all but disappeared). There was a much closer relationship between sounds and letters than exists today, but notice that there was no sense of a correct way to spell: the outmoded pronouns *thow* (= thou) and *þyne/thy/þi* (= thine/thy) fluctuate between the old Runic symbol <þ> and French-inspired <th> to represent [ð]; <y> and <i> are both used.

There is also plenty of odd-looking grammar in these three lines. The form *etist* '(you) eat' shows an obsolete verb ending *-est* (briefly revived by the internet phenomenon LOL-speak (a curious interpretation of how cats might speak English): *Canst thine hath milk curd and minced beef sandwich?*). The negative construction is no longer grammatical either. In Modern English, we need the support of the auxiliary *do* as in *Do not belch*. Although *do* was available to speakers at this time, we waited 400 years for it to became obligatory in negation. (The text has *bulk*, an obsolete verb meaning 'to belch'; it is probably a relative of *belch*, though this is not entirely clear.) The form *were* in *as a Beene were yn þi throte* shows a grammatical feature known as the subjunctive. This was a handy verb marker to signal unreality; even at this time, its function was being taken over by verbs such as *may, might, will* and *would* (for example, *as if a bean might be in your throat*). Finally, there is also the now archaic form *whylis þat*, the source of our modern conjunction *while*. Here we see the noun form *while* 'period of time' with the *-s* ending once used to make adverbs (compare *nights*, as in *work nights* meaning 'by night').

Any investigation of change in language has three obvious components:

- the 'what' of change – possible and impossible changes
- the 'how' of change – processes and mechanisms of change
- the 'why' of change – explanations for change.

We have already seen in Chapters 4 and 6 the extent to which vocabulary can change – words come and go and meanings shift constantly. In this chapter, we move our attention to the types of changes that alter the structure of languages, and consider the above three components as they relate to phonology and grammar. We also explore the linguistic relationships that arise as a consequence of change, and show how linguists determine genetic relatedness and group languages into their families.

11.1 The ‘what’ of change: Phonology

> It would not be entirely inappropriate to regard languages in their diachronic aspects as gigantic expression-compacting machines.
>
> Ronald W. Langacker (1977), *Syntactic Reanalysis*

In Chapter 8, you were introduced to a number of phonological processes. You saw how reductions and assimilations (for example, *miss you* [mɪʃə] and *did you* [dɪdʒə]) are part and parcel of ordinary connected speech – fast-speech phenomena often with no long-term effect. However, reductive processes can leave an enduring impression on the language. Precisely the same process that produces [mɪʃə] and [dɪdʒə] gave English its palatal phoneme /tʃ/ in words such as *chin* and *cheese* (Old English *cin* [kɪn] and *cese* [kesə]). These kinds of historical changes are indicated with an arrowhead to denote the direction of the change ([kɪn] > [tʃɪn]), to distinguish them from processes that are underway in the language, shown with a complete arrow ([mɪsjʉ] → [mɪʃə]). Also note that we use the asterisk in this chapter to indicate a form that has been reconstructed from more recent forms – for example, **samftō* is reconstructed from Old English *sōfte*, Modern English *soft* and Modern German *sanft*. (This is in addition to the use of the asterisk to indicate ungrammaticality.) We explain the process of reconstruction in more detail in section 11.5.

Old sounds drop out: Deletion

Time sees the **deletion** of consonants, vowels and even whole syllables. In the case of consonants, loss is sometimes balanced by a lengthening of the neighbouring vowel; the vowel length becomes a phantom of the lost consonant. This is called **compensatory lengthening**. Consider the examples in Table 11.1.

Table 11.1 Loss of nasals in a selection of English words

Proto-Germanic	Old English	Modern English	Modern German
**samftō*	*sōfte*	*soft*	*sanft*
**fimf*	*fīf*	*five*	*fünf*
**munþ*	*mūþ*	*mouth*	*Mund*
**tanþs*	*tōþ*	*tooth*	*Zahn*
**gans*	*gōs*	*goose*	*Gans*
**uns*	*ūs*	*us*	*uns*

Old English shows a sound change that affected the sequence vowel + nasal + fricative, as shown in the parent Proto-Germanic. In each instance, the loss of the nasal was accompanied by a lengthening of the vowel (signalled by a dash). The change didn't affect German; it preserves the original nasal (don't worry about the various additional changes that have occurred – they have other stories to tell).

Change that involves the deletion of entire syllables is called **haplology**. Typically, it involves consecutive syllables that are similar-sounding, and the history of English has a number of well-known examples, the most famous being the name *England* itself (< *Anglaland*, meaning 'Land of the Angles', one of the original Germanic tribes that settled in Britain). The word *eighteen* once used to have four syllables (< Old English *eahtatyne*) and *eighty* had three (< Old English *eahtatig*); both words have undergone many changes, one of which involved the loss of the middle [ta] syllable. Both German and English show regular 'haplogy' when the addition of an affix creates successive similar-sounding syllables – for example, *Zauberin* 'sorceress' (< **Zauber-er-in*) and *sorceress* (< **sorcer-er-ess*); *gently* (< **gentle-ly*). In early Japanese, syllables were regularly deleted if they involved sequences of identical consonants except in voice – for example, *kupibisu* 'a heel' became *kubisu.*

New sounds appear: Epenthesis

Segments can also appear where they previously didn't exist; like history repeating itself, *England* is sometimes pronounced as [ɪŋgələnd]. The process, known as **epenthesis**, is extremely common, especially when difficult consonant clusters are involved. We saw, in Chapter 8, how Japanese shows regular vowel epenthesis in *gairaigo* 'loanwords' such as *resutoran* 'restaurant'. While epenthesis is generally less common than deletion, it can result in long-term change. Romance languages show the regular addition of a vowel at the start of words beginning with /s/ and another consonant – for instance, Latin *status* 'state' > Spanish/Portuguese *estado* and Old French *estat* but Modern French *état* (note that both forms, with and without the epenthetic vowel, came into English as *estate* and *state*).

In Chapter 8, we also saw how consonants are sometimes added as the vocal organs move from one position to another. In Japanese, the consonant /r/ is inserted epenthetically between a vowel final verb stem and a vowel initial suffix: compare vowel final *mi-ro* 'see' (imperative) with consonant final *mak-e* 'scatter (imperative)'. Table 11.2 shows examples of epenthesis in English. Where possible, we have supplied some German forms that maintain the Proto-Germanic consonants – that is, with no epenthesis. (Incidentally, [p] was added to *empty* and not [b] (as in *bramble*) to reflect the voiceless quality of the following [t].)

Table 11.2 Loss of nasals in a selection of English words

Old/Middle English	Modern English
brēmel	*bramble*
[illegible]	*mumble*
[illegible]	*crumble*
gandra (earlier form *ganra*)	*gander* (compare *gannet*)
þunor	*thunder*
ǣmtig	*empty*

Recall that mistiming can also cause consonants to appear at the ends of words. In Table 11.3, the German words show the epenthetic [t], and English is the more conservative this time. Enhancing word endings with extra sounds was once common in English, too. Both *twice* and *once* had *t-* versions – *wonst* and *twyst* – as did *nice, close, across* and *since*. Some *t-* afterthoughts didn't drop off: *against, peasant, truant* and *parchment* are now standard. Occasionally, words retain alternative pronunciations – interestingly, original mispronunciations such as *amongst, amidst* and *whilst* now sound literary (even highbrow!) compared with *among, amid* and *while*.

Table 11.3 Final consonant epenthesis in German

Modern German	Early English
Habicht	*hawk*
Axt	[illegible]
Sekt ‘(bubbly) wine’	*sack (sherry)*
Palast	*palace*
[illegible]	[illegible]
Papst	*pope*
selbst	*self*
einst	*once*

Phonotactics, shmonotactics

Recall that the arrangement of sounds within words is known as phonotactics. It's all about what sort of consonant clusters and vowel sequences a language allows, and this can change through loanwords. English has been taking on many Yiddish expressions and these are introducing new combinations at syllable onset: *shmooze* ‘to fawn over someone for personal benefit’; *shlep* ‘to carry, lug’; *shlock* ‘inferior merchandise, low-grade entertainment’; *shlemiel* ‘awkward unlucky person’; *shmuck* ‘idiot’; *shmaltz* ‘excessive sentimentality’; *schnozz* ‘the nose’; *schnook* ‘an easily deceived person’; and *schnecke* ‘sweet roll’ (the type we generally know by the French name *escargot*). The spelling is flexible – both < sch > and < sh > are used. Up until the arrival of these loans, the sound we represent as [ʃ] was rather restricted with respect to the company it could keep. German loanwords had given us the [ʃv] cluster, as in *schwarmerei* ‘unwholesome sentiment’, and [ʃw], as in *schwa* ‘a type of vowel sound’ (the pronunciation of these words depends on how well integrated these German words are). Yiddish borrowings are now adding three additional consonant combinations: [ʃl], [ʃm] and [ʃn].

Existing sounds are modified: Assimilation, dissimilation and metathesis

As we outlined in Chapter 8, sounds can change their shape to become more like neighbouring sounds (**assimilation**) or less like them (**dissimilation**); they might even swap positions with them (**metathesis**). Not only are these processes omnipresent in ordinary connected speech, they are also the source of common types of sound change.

Most changes are assimilatory, and this can involve many different types of accommodation. Recall how nasals (the chameleons of the phonological world) are forever altering their place of articulation to blend into their environment (compare *indecent* with *impossible,* where /n/ conforms to the place of articulation of the following consonant). In multi-syllable words in Japanese, there is also a strong pattern of nasal sounds in syllable final position taking on the place of articulation of the following consonant. In fact, the *kana* orthography has just one representation for a nasal coda, relying on the form of the following syllable to inform readers about the place of articulation for the nasal. Thus, the nasal in the word *san* ‘three’ is velar in [saŋgatsu] (‘March (third month)’), bilabial in [samban] (‘third’), and palatal in [saɲju] (‘thirty’).

Palatalisation is one of the most common types of assimilation. In English, it is responsible for the current pronunciation of words such as *nation, sure, sugar, measure* and *pleasure* (spelling occasionally gives a clue here – for instance, the <ti> in *nation* shows the original [i] that triggered the palatal [ʃ]). It is even possible for sounds to assimilate totally to neighbouring sounds (that is, across all parameters: place of articulation, manner of articulation and voicing). Consider the history of the word *wool.* The Proto-Germanic cluster [ln] became a long (or geminate) [ll] in early English and German (now pronounced [l] in both languages).

(11.2)

Proto-Germanic **wulno* > **wullo* >

Old English *wull*/Old High German *wolla*

Occasionally, non-contiguous sounds assimilate. In consonants, the process is unusual, responsible for sporadic changes like *orangutan* > *orangutang* (an example repeated in all textbooks because the process is so rare!). Distance assimilation of vowels, **vowel harmony**, is more common. English used to harmonise certain vowels and relics of this remain in irregular plural nouns (for example, *foot–feet* and *mouse–mice*). These forms are left over from an earlier pronunciation rule whereby the main vowel of the word fronted to be more like the vowel of the plural *-i* ending. This can be illustrated in reconstructed forms: **fōti* 'feet' and **mūsi* 'mice' (< Proto-Germanic **fōtiwiz* and **mūsiz*). When the ending eventually dropped off, we were left with a strange-looking alternation that, through later changes, eventually became modern *foot–feet* and *mouse–mice*. Table 11.4 reconstructs how it would have worked (note that [ø] is a fronted [o]).

Table 11.4 Vowel harmony in early English

Singular	**Plural**
[fo:t]	[fo:ti] (in Pre-Old English)
[fo:t]	becomes [fø:ti]
[fo:t]	becomes [fe:t]
[fut]	becomes [fi:t] (in Modern English)

This kind of vowel harmony has resulted in many other irregular vowel alternations that are still present in Modern English – for example, verbs pairs such as *full–fill, food–feed, lie–lay, fall–fell* and *drink–drench*, and adjective sets such as *old–elder–eldest* are relics of this process.

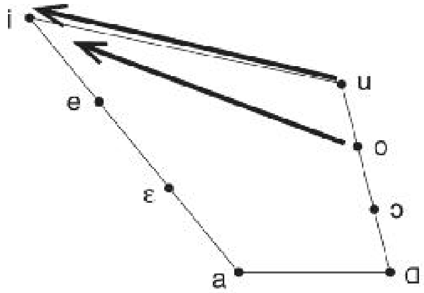


Figure 11.1 The fronting of back vowels

So why isn't *hice* the plural of *house*?

The Old English word for *house* was *hūs*, so you might imagine it would pattern like *mouse* (Old English *mūs*). However, *house* originally had no plural ending; it patterned like *one deer – two deer*. Only later, in the twelfth century, was it regularised to *husas* or *huses*, and from the fourteenth century onwards to *houses*. So, there wasn't an ending to trigger any vowel change, and *hice* was never an option. It is not likely to be introduced either, since the *mouse–mice* pattern isn't lively in English. But, speakers do love to play with language, and occasionally irregular patterns get a new lease of life. A jokey plural of *spouse* is now *spice*, inspired by the *mouse–mice* pattern, but with the extra piquancy that *spice* brings.

We saw in Chapter 8 how the process of vowel harmony is still active in Modern German to the extent that the vowel change (known as *umlaut*) has its own marker (or diacritic). German orthography uses two dots above a vowel to indicate it is a fronted version of the normal vowel letter (so and <ü> represent distinct sounds). Vowels written as <a>, <o>, and <au> become <ä>, <ö>, <ü> and <äu>; take, for example, plural forms such as *Mann* 'man' – *Männer* 'men'; *Fuß* 'foot' – *Füße* 'feet'; *Haus* 'house' – *Häuser* 'houses'. The *umlaut* illustrates one of the great puzzles for historians of language: how (and why) do close siblings end up diverging in this way; in this case, why did the *umlaut* bite the dust in English but endure in German?

The *umlaut* – the 'coolest of all diacritical marks'

In the 1970s, there appeared a more decorative use of this diacritic (the 'metal umlaut' as it is known) in the names of heavy metal bands like Motörhead, Queensrÿche, Mötley Crüe – and wonderfully spoofed in the parody metal band Spin¨al Tap. The umlauts here don't affect the pronunciation of these band names in any way, but communicate (especially when combined with Gothic script) strength and aggression, traits stereotypically associated with Germanic cultures – think Gothic horror, marauding Vikings. During an interview for *The Wave* magazine, Lemmy Kilminster, founder of Motörhead, described the umlaut this way: 'I only put it in there to look mean'. Following the days of heavy metal, many other groups have taken to stylising their names with umlauts and creative capitalisation. These include Toilet Böys, Jaÿ-Z, dälek (pronounced 'die-a-leck') – even Lady

Gaga's song 'Yoü and I' comes with ümlaut garnish.

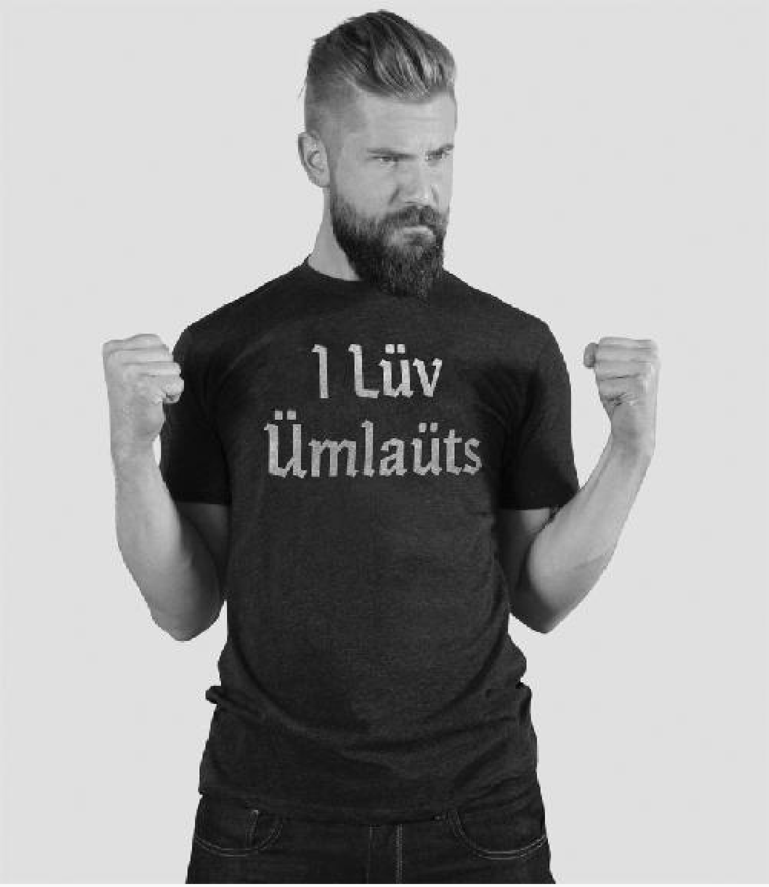


Figure 11.2 Headline t-shirt celebrating *umlauts*

Finally, when sounds devoice (lose their voiced quality) at the ends of words, this is a kind of assimilatory change; think of the voiced consonants as assimilating to the voicelessness of the following silence. You have already seen German examples of this in Chapter 5 (Table 5.13), but we give others in Table 11.5.

Table 11.5 Final consonant devoicing in German

Old High German	**Modern German**	**Transcription**	**Gloss**
tag	*Tag*	[taːk]	'day'
rad	*Rad*	[ʁaːt]	'wheel'
hunt	*Hund*	[hʊnt]	'dog' (compare 'hound')
bad	*Bad*	[baːt]	'bath'

Far less common than assimilation is dissimilation, and although nasals and liquids ([l] and [r]) are prone to the process, it rarely occurs on a regular basis. In Table 11.6, we have bolded the consonants that have dissimilated.

Table 11.6 Examples of dissimilation in English and German (changed consonants bolded)

Early form	Modern English	German
peregrin (from Latin)	*Pi**l**grim* (compare *peregrine falcon*)	
turtur (from Latin)	*turt**l**e*	
purpura (from Latin)	*purp**l**e* (compare the heraldry term *purpure*)	
marmor (from Latin)	*mar**b**le*	
**eb(e)n-*	*even*	*eben*
**hemel-*	*heaven*	*Himmel*
cardamom	*cardamon* (non-standard)	

You will notice that, like assimilation, the process of dissimilation can involve different features, and can also work in different directions. In a sequence of [r … r], English (and occasionally German) changed one [r] to an [l]. The first nasal in a sequence was changed to a non-nasal sound; in *even* and *heaven,* this was a voiced fricative, which in German (*eben*) became a stop, while the German *Himmel* underwent dissimilation of the second nasal. In the case of *cardamom,* some Modern English speakers (including both the current authors!) dissimilate the final [m].

Another example of dissimilation can be found in Japanese. Many sequences in Japanese are subject to Lyman's Law, which states that only one voiced stop, fricative or affricate can occur within a morpheme. Normally, the first consonant in the second part of a compound attracts voicing (for example *ori* 'fold' + *kami* 'paper' gives us *origami* 'the art of paper folding'). However, when the second element contains another voiced stop or affricate, this voicing change is blocked (compare *ao* 'blue' + *same* 'shark', which takes on voicing *ao zame* 'blue shark' with *ao* 'blue + *sagi* 'heron', which does not because of the /g/: *ao sagi* 'blue heron').

Another unusual process is metathesis (or 'methatesis'), whereby adjacent segments reorder themselves (often they involve the notoriously unstable nasals and liquids). It starts off as a kind of mistiming that can remain a one-off slip of the tongue (or childhood error), but when a number of speakers keep making the same adjustment, tongue slips can endure. Table 11.7 provides some examples in English, with German equivalents where possible (sometimes showing metathesis, sometimes not). We are sure you can think of many more!

Table 11.7 Examples of metathesis in English and German

Early English	Modern English
þridda (compare *three*)	*third*
brinnan	*burn*
beorht	*bright*
brid	*bird*
crud	*curd* ‘coagulated milk’ *crud* ‘dirt’
wæps	*Wespe*
[illegible]	[illegible]
Anemone	*Amenone* (non-standard)
taske	*task* ‘mission’; *tax* ‘toll’

In the discussion of syllables in Chapter 8, we saw how, around the world, there are secret languages that draw on syllable structure to distort words in different ways. In Japan, metathesis of whole syllables was an important strategy in the development of a secret language used by thieves between the Edo and Showa periods. Examples include *kabe* ‘wall’ becoming *beka, tabi* ‘journey’ becoming *bita* and *mane* ‘money’ becoming *nema.*

Phonetic versus phonological change

In Chapters 7 and 8, we made a distinction between a phonetic description (the physical aspects of sounds) and a phonological description (how sounds relate to each other to form a system). The history of languages can show both types of changes.

Phonetic changes involve adjustments to the realisation of a phoneme or to the distribution of the allophones of a phoneme, but cause no disruption to the number of phonemes or the relationships between them. For example, the phoneme /r/ is now usually an approximant but was once pronounced as a flap or trill, as in Scots English today. So, the shape of the sound changed acoustically, but this didn’t alter the phonology of English.

We also know from spelling and other sources (such as modern dialects) that at one time speakers of English articulated /r/ wherever it appeared in a word – for instance, *father* was pronounced something like [faðər]. Sometime between 1600 and 1700, however, the consonant disappeared except where it appeared before vowels. Since the English colonisation of America took place earlier, American English remained rhotic or ‘R-ful’. Both conservative ‘R-ful’ and innovative ‘R-less’ dialects have an /r/ phoneme, but its distribution has changed:

(11.3) /r/ > [ɹ] before vowels
ø everywhere else

The ears and brains of Australian and New Zealand English speakers continue to ‘hear’ a rhotic at the end of *father*, and it resurfaces when a vowel follows, as in *father and mother* [fɐːðəɹənmɐðə]. These sorts of phonetic adjustments are going on all the time in languages, although past examples are often difficult to determine because they don’t tend to be reflected in the writing system (and this becomes increasingly the case as spelling is conventionalised).

We now turn our attention to an example of the more significant phonological change that impacts the whole system. To illustrate, we take the evolution of a new nasal consonant in English and German. Both languages currently have three nasal phonemes:

(11.4)

Bilabial	Dental	Velar
m	n	ŋ

In Chapter 8, we described the velar /ŋ/ as a new phoneme that doesn’t show the same healthy distribution as /m/ and /n/ – it can only occur word finally and between vowels (which is why speakers of English and German have trouble with surnames like *Ng* and *Nguyen*).

On the basis of evidence from dialect variation, spelling, rhyme, early descriptions and our knowledge of sound change, we know this phoneme began life as a pronunciation variant of /n/, occurring only before other velar sounds (remember the pronunciation of *pancake* as [pæŋkæɪk], evidence of the chameleon-like nature of the nasal). In other words, [ŋ] was simply an allophone of /n/. So, an English word like *sing* was pronounced [sɪŋg]. The pronunciation rule was something like this:

Early English and German pronunciation of /n/

(11.5)

/n/	[ŋ]	before {k, g}
	[n]	everywhere else

How, then, did the velar nasal acquire its new status as a phoneme? At some stage during the history of both languages, stops fell off the end of [nasal+stop] clusters – hence the silent <b> in *lamb* and *climb*, and the modern standard pronunciation of *sing* and *long* with final [ŋ] not [ŋg]. English regional pronunciations of *sing* as [sɪŋg] are relics of this earlier pronunciation, as are German dialect pronunciations of *lang* 'long' and *jung* 'young' with a final [k] (compare the relic English pronunciations *nothink* and *everythink*).

Once [g] disappeared from words like *sing*, the velar nasal came to be in contrastive distribution with the other nasals in this position – *some, sun* and *sung* now make a minimal group for English; the meaning contrast between these words depends entirely on the distinction between the nasals.

Fingers don't fing

Why do we pronounce *g* in *finger* but not *singer*? In early English, the loss of [b], [d], [g] after nasals was only ever at the end of words. Sounds that were stuck in the middle were protected, as in *finger, angry, hunger* and *longer*. So, a word like *singing* (which we now pronounce as [siŋiŋ]) would have been pronounced [siŋgiŋ]. The [g] in the middle was safe.

During the 1700s, people started to expand the 'g-less' pronunciation to related words. So, [g] not only disappeared from *sing*, but also forms that derived from *sing* – namely, *singing* and *singer*. But, since a *finger* is not something that 'fings', the [g] in *finger* was preserved. (In fact, historically *finger* is a derived form, being related to *fang*, but that was just too long ago to be relevant!) What we are seeing here is **analogy** at work. The 'g'-less pronunciation was such a hit that speakers extended it to other words. Besides, it was messy that *sing* was pronounced differently from related *singer* and *singing*.

Okay, you might be wondering how come we still pronounce [g] in derived forms, such as *younger* and *longer* (from *young* and *long*)? Other adjectives, such as *cunning*, have lost their [g] (pronounce *he's the cunningest chap I know*). And, if we coined a new adjective, say *tring* 'attractively miniature', the derived forms *tringer* and *tringest* would also be ‘G-less’. As we saw with peculiar nouns like *foot–feet*, common usage words are often exceptionally conservative. Their frequency gives them a boost and reinforces their irregularity, so they haven’t yet fallen in line.

The acquisition of the new velar nasal had the result of introducing some balance into the sound systems of English and German. Oral stops are typically matched with nasal stops at the place of articulation and, as the partial English consonant chart shows below, the lack of a velar nasal was an obvious gap.

Old English **Middle English**
p t tʃ k
b d dʒ g
m n ŋ

While we can’t go as far as to say that all sound change brings about a more balanced sound system, many changes in English and German have had the effect of introducing greater symmetry – gaps were either filled or eliminated somehow. But, this is only a tendency and it remains to be seen whether English acquires a palatal nasal phoneme to match [tʃ]/[dʒ].

11.2 The ‘what’ of change: Grammar

> Grammars code best what speakers do most.
>
> John W. Du Bois (1985), Competing Motivations

Grammar – and especially syntax – has been a rather neglected area in historical linguistics. While some headway was made early on in understanding changes in word structures, it wasn’t until recently that much was known about syntactic change. One of the reasons is that grammar is just so much harder to investigate. On a practical level, we need much more data. And, unlike phonological inventories (which are finite), there are thousands of different grammatical constructions, and hundreds of different grammatical categories.

Change in morphology: Morphological types

> A linguist that insists on talking about the Latin type of morphology as though it were necessarily the high-water mark of linguistic development, is like the zoologist that sees in the organic world a huge conspiracy to evolve the race-horse or the Jersey cow.
>
> Edward Sapir (1921), *Language*

When you were introduced to our case-study languages, you would have seen that they differed a lot with respect to how much meaning could be packed into a single word. In this regard, languages fall into three main groups; over time they can shift from one type of morphological structure to another.

Isolating (analytic): Turung, Chinese, Vietnamese

Though most languages fit a type to a matter of degree, a pure **isolating** language would have no bound morphemes at all – words would consist of single morphemes only, with the grammatical function of words signalled by word order or by free-standing forms. Our case-study language Turung is isolating to a large degree (see example 11.6).

(11.6)

[illegible]
[illegible]
‘then (they) would tie thread’

Agglutinating: Japanese, Turkish, Swahili

The term **agglutinating** comes from the Latin word *gluten* (‘glue’) and it covers languages with words that are easily dissected into sequences of morphs. As example 11.7 (in Japanese) illustrates, it is as if the morphs were glued together – one morph represents one morpheme and the boundaries between them are clear-cut.

(11.7)

Jirō-ga tegami-o yom-u
Jirō-subject letter-object read-nonpast
‘Jirō reads a/the letter(s)./Jirō will read a/the letter(s).’

In these languages, affixes are used over and over again to build new words. They retain the same shape, except for occasional changes resulting from regular phonological rules of the language. Most of the world’s agglutinating languages are suffixing – for example, Japanese and Turkish. Swahili is an odd one out (with prefixes).

Inflectional: German, Russian, Sanskrit

These languages can have many morphemes within a single word, but the boundaries between the morphs are inconsistent and there is a lot of tricky allomorphy. Consider the following example from German: *Tias sieht Fritz* ‘Tias sees Fritz’. The verb is *sehen* ‘to see’, and the inflected form *sieht* [zi:t] contains the following meanings and features:

1 the idea of seeing

2 event takes place in the present

3 subject is the third person

4 subject is in the singular.

In a prototypical agglutinating language, each of these four meanings would have clearly identifiable parts, but in *sieht* we can’t draw clear boundaries in this way. Moreover, change any of these meanings and the verb takes a different form. In the sentence *Ich sah Fritz* ‘I saw Fritz’, the verb *sah* [za:] contains the following four meanings and features:

1 the idea of seeing

2 event takes place in the past

3 subject is the first person

4 subject is in the singular.

Some linguists refer to fused (unsegmentable) words like *sah* as ‘portmanteau morphs’ (a portmanteau is a bag with multiple compartments). There is no one-to-one correspondence between the various categories (person, number, tense and so on) and their formal marking – everything is packed together in the one portmanteau!

In Chapter 4, we introduced you to a fourth (rarer) type of morphological structure: the **polysynthetic** languages such as West Greenlandic. The term ‘synthetic’ describes languages with words that usually contain more than one morpheme (and so includes both inflectional and agglutinating). These languages are extreme forms of synthetic languages and are notorious for their complex words. Example 11.8 is from an Australian Aboriginal language, Tiwi. This one word is rendered into English with a sentence, ‘She is walking over there in the morning with a light’. The gloss provided underneath indicates the morpheme boundaries (as before, don’t worry about the grammatical labels here – just admire the morphological complexity).

(11.8) *ampiniwatuwujingimajirranirningiyangurlimayami*
a-mpi-ni-watu-wujingu-ma-jirrakirningi-angurlimayi-ami
3 minS.F-npst-loc-morning-dur-com-light-walk-ipfv
‘She is walking over there in the morning with a light’

This Tiwi word has nine different morphemes; some of them would be treated in other languages as independent words (such as *light* and *morning*), yet they are integrated into one single word.

We can quantify word complexity in languages by calculating the average number of morphemes per word – the lower the ratio, the more isolating the language. A ‘pure’ isolating language would have a ratio of 1.00 (so a one-to-one relation of morphemes to words). Modern English, at 1.68, is much more isolating than Old English, at 2.12, or Sanskrit, at 2.59. The ratio for polysynthetic Greenlandic is a whopping 3.72!

For linguists of the nineteenth century (such as the von Schlegel brothers, who led the way here), the three main morphological types represented different stages in the evolution of language. Like others of their time, they were strongly influenced by Darwin, believing that each language had a life-cycle and, like plants and animals, was subject to natural evolution. The inflectional stage (exemplified by classical languages such as Greek, Latin and Sanskrit) was regarded as the highest linguistic level a language could attain – loss of morphological complexity was viewed as decay, and predominantly isolating languages like poor old English were felt to have fallen on hard times.

While these linguists were mistaken in attributing values to these stages, their idea of a morphological cycle has support. As you will see in a moment, there are changes that drive isolating languages towards agglutination, there are changes that drive agglutinating languages towards inflection and there are certainly changes that drive inflecting languages towards a more isolating structure. But, we can't say that any one of these stages is linguistically superior to another – just as zoologists can't say racehorses are better than quaggas!

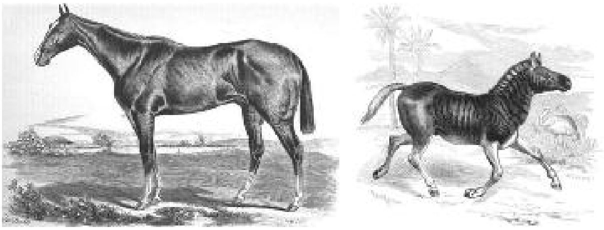


Figure 11.3 The horse and the quagga

Change in morphology: Analogy

Analogy involves speakers regularising linguistic patterns so that forms match up with other forms that speakers see as similar in some way. There are many types, but all exploit this same principle of resemblance to an existing pattern. You have seen analogy at work in the lexicon (in metaphor and in the extension of productive affixes to create new words like *bookoholic*) and in phonology (the extension of the 'G-less' pronunciation from *sing* to derived forms like *singer*). Reasoning from analogy motivates many different morphological changes.

Consider the extension of the *-ed* past tense ending to the new(ish) verb *to twerk*. We can express it in terms of a proportional equation:

(11.9) *work : worked* as *twerk* : ?
? = *twerked*

To see how analogy brings about morphological change, consider the verbs in Table 11.8.

Table 11.8 Analogy at work in English verbs

[illegible]
[illegible]

If you extend the *-ed* pattern to *go*, you arrive at *goed*. We all understand children or language learners who say *goed*, but it has not yet become standard. On the other hand, *shined* and *shone* coexist (though *shined* is overtaking *shone*), and *climbed* has well and truly knocked off *clomb*. For patterns to generalise in this way, they have to be lively (or 'productive'). Consider the verbs (all sharing the basic rhyme [ɑe]): in Table 11.9.

Table 11.9 Basic and past forms of [ɑet] verbs in English

Basic	**Past**
write	wrote
fight	fought
cite	cited

Only *cited* represents a productive process of past tense formation – it's the one that will extend into new situations. We can understand the language learner who says *I writed* or *I fighted*, but not the one who comes out with *I wrought, I cought, I fote* or *I cote!* These marginal past tense patterns do not have the capacity to extend by analogy. Bear in mind, though, that this is language and occasional unproductive patterns take off in surprising ways (recall the playful plural *spice* from *spouse!*).

Analogy has often been described as a cleaning-up process that steps in to tidy up the mess left by sound change. We have seen, for example, how sounds will often drop out or assimilate with their neighbours and this can introduce complications into the grammar. Recall those badly behaved nouns like *leaf-leaves* and *wife-wives* we mentioned in Chapter 5 (the discussion of allomorphs). These forms show a voiceless consonant in their singular forms but a voiced consonant in their plural forms – the residue of an old pronunciation rule that saw voiceless fricatives like /f/ and /s/ become voiced when they were flanked by voiced sounds like vowels – a simple enough change (assimilation to the voiced vowel sounds on either side).

This pronunciation rule eventually stopped being a living part of the language and when this happened it left behind a group of irregular older nouns – *leaf-leaves* versus regular *brief-briefs*. Over the past few hundred years, these forms have gradually been brought into line via analogy; the plural of *cliff*, for example, is no longer *clives*. New nouns will automatically step into line. Take a fairly new noun snorf, 'to describe a quick drink of liquor'. Its plural is definitely *snorfs* not *snorves*. True, *lifes* and *leafs* haven't caught on yet, but *still lifes* sounds good, and so does the name the Canadian *Maple Leafs*.

Table 11.10 Analogy at work in English nouns

Two words for the price of one

Very occasionally, analogy can mop up in a different direction. Consider what happened to the noun *glove.* The singular form was originally Old English *glōf,* ending in a voiceless fricative; its plural would have been formed with the voiced fricative [v] (giving the Modern English *gloves*). Normally, the outcome of analogy would have been *glof* and *glofs*. Instead, the modern singular form *glove* has been created from the plural *gloves* – the regularising has worked in the opposite direction. Presumably, the fact that gloves usually come in twos meant that the plural form was more common and that's why we find this unusual development. But, analogy sometimes does do surprising things. For example, it has even worked in both directions, giving rise to two completely different words. This is what happened to the word *staff* 'stick, rod' (Old English *stæf*) – the modern language has both *staff–staffs* 'stick' and *stave–staves* 'line in musical notation'. A similar split happened to the plural of the word *cloth* 'material' – *cloths* versus *clothes* 'garments' (Old English *claþ*).

Change in syntax: Word order types

In Chapter 9, we introduced an important classification based on the order of major constituents: S(ubject), V(erb), O(bject). All three major word order types (SVO, SOV and VSO) are found in our different case-study languages, but it is also possible for languages to change type.

English and German once had considerable flexibility in word order – in fact, all logically possible arrangements of subject, verb and object were possible. Examples 11.10–11.14 from Old English (449–1100 ce) give a hint of the possible variation at the time. To make things straightforward, we have considered only the ordering of the subject (=S), the main verb (=V) and everything else (=X) in main clauses. The sentences come from a medical account of someone who was *ælfsogoþa* ('elf-sucked'), a nasty affliction caused by elf-possession (Anglo-Saxon texts are full of descriptions of assaults by supernatural beings and sprits of evil; in this case, the condition was probably anaemia).

Examples 11.10 and 11.11 both show main clause verbs in second position (SVX and XVS respectively); verb second was the most usual order.

(11.10)

[illegible] Verb-second
[illegible] sogaþa
[illegible] sucked
‘He is possessed by elves’

As an interesting aside, example 11.10 has a curious subject form, *him*. In early Germanic, speakers could indicate that entities weren’t wilfully involved in an event by placing them in the dative case (that is, *him* rather than *he*). German, with its case system intact, still has this construction, but English only has relics (*methinks*).

(11.11)

[illegible] Verb-second
[illegible]
[illegible]
‘Then you can improve it for a while’

The next examples show three rarer orders: verb-third (XSV) order (see example 11.12); the expressive verb-first (VSX) order more typical of vivid prose (see example 11.13); and the linking verb-final or near-to-verb-final order (SXV(X)) order that was common in clauses in sequence (see example 11.14).

(11.12)

[illegible] Verb-third
[illegible]
[illegible] completely

‘That man you might cure completely’

(11.13)

[illegible] Verb-first
[illegible]
[illegible]
‘They will immediately disappear’

(11.14)

[illegible] Verb-final
[illegible]
[illegible]
‘He shall end his life by an evil death’

As the Modern English versions show, ordering is now rigidly SVX. Modern German also has fixed word order, although it is different typologically from English, since it has preserved some of the original verb-final tendencies (for example, in subordinate clauses). We haven’t got the space to go into why these two languages abandoned this earlier flexibility in favour of a system of grammatically rigid ordering, but rest assured that there were undoubtedly many different interlocking factors involved here. As we discuss later in this chapter, language change is never one-dimensional and the explanations for it aren’t either.

Change in syntax: Creating grammar

Have you ever wondered where grammar comes from? What’s behind the *-ed* ending on *jumped*, the plural ending on *books*, the future marker *will* in *I will jump* and so on? Looking into the history of grammatical bits and pieces, we find they have often sprouted from ordinary words – everyday lexical items (those that relate to aspects of human experience). The creation of grammar is called

grammaticalisation. Linguists studying it examine the lexical sources that give rise to grammatical forms and constructions, as well as the changes that take place as items become more grammatical.

Here we discuss a new future marker in both English and Pennsylvania German; it has evolved from the lexical verb of movement *go/geh*.

(11.15) Fred's *going* to university. (lexical)

(11.16) Fred's *going to/gonna* go to university. (grammatical)

Example 11.15 illustrates the lexical sense of *go* (Fred is physically on a path moving towards a goal) and example 11.16 is the more recent abstract sense (Fred is on a course towards a particular end point in the future.) When grammar is created in this way, it involves processes at all different levels of the language. Here we see a bleaching of the original lexical meaning – the new future marker *going to/gonna* shows a shift from movement in space to movement in time.

Accompanying the loss of lexical status is also a reduction of vowels and consonants. What was originally three morphemes (*go* + *ing* + *to*) is now an unsegmentable whole, ***gonna***. As items reduce, they may become more dependent on surrounding material, perhaps even fusing with other words in the vicinity (for example, *will* > *'ll*, *not* > *n't*, *him/them* > *'em*, *is* > *'s*). This gradual erosion gives rise to the scenario in Figure 11.4: words > clitics > affixes.

words > clitics > affixes

Figure 11.4 Phonological reduction of words

The meaning of the source element(s) often shapes the subsequent grammatical meaning. *Go*-derived futures are common among the world's languages, and they are typically 'immediate futures'. Example 11.17, from Pennsylvania German, captures the sense of immediacy that is typical of these future markers.

(11.17)

[illegible]
[illegible]
‘I thought – it’s gonna happen to him (at any moment)!’

The immediacy of *geh* (and *gonna*) is a fallout of the lexical source (a movement verb) – in other words, the idea of something already being on a path and moving towards some future goal. You can see this more clearly if you compare the three different English future markers in example 11.18.

(11.18) The ladder shall/will/is gonna fall.

With *the ladder’s gonna fall* you can almost see the ladder tumbling!

And I’m like …

The word *like* is among the most versatile words in English – it can be an adjective (*like opinions*), preposition (*he’s like a god*), subordinator (*tell it like it really is*), adverb (*anything like so well done*) and, more recently, discourse marker for focus (*I’m not like depressed, just a bit grumpy*) or for hedging (*Do you wanna like see a movie?*). Of all these roles, the most controversial is still the ‘quotative *like*’, which introduces reported speech, thoughts or attitudes, facial expressions, even noises (*I looked at him and I’m like gimme a break will ya*). All are spectacular examples of grammaticalisation at work.

All these *likes* have their origins in the Old English noun *lic* ‘body, form’ (preserved in the modern *lychgate* ‘the roofed gateway to a church yard’ – see Figure 11.5). In fact, Old English *līc* is also the source of the suffix *-like* (as in *godlike*) and even the *-ly* suffix (as in *godly*). So, *godlike* and *godly* have the same origin – a nice piece of linguistic history repeating itself.

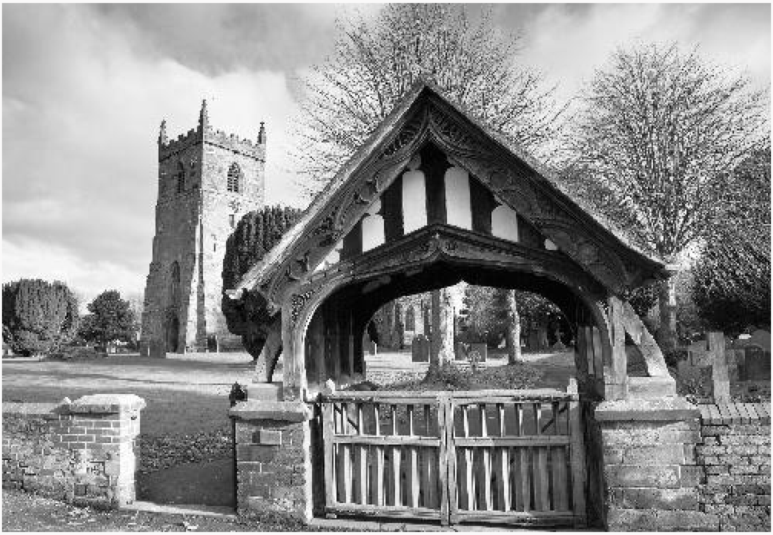


Figure 11.5 A church lychgate

Are all lexical items eligible for grammaticalisation? Could a word like *pavlova* ever evolve into an indirect object marker? Probably not! The lexical items that are pressed into grammatical service are generally drawn from the most everyday basic vocabulary. They are culturally independent (unlike *pavlovas*); they refer to the most fundamental of human experiences and, most importantly, they are used frequently. We return to this aspect in a moment.

11.3 The 'how' of change

New expressions, even new pronunciations and structures, can take off because of faddish demands or because of the influence of certain speakers. Often it is celebrity endorsement that propels a linguistic feature to success. The verb *to dog* didn't take off until Shakespeare used it; and the verb *to twerk* got a huge boost when Miley Cyrus 'twerked into history'. There was a spate of discussion about *conscious uncoupling* after Gwyneth Paltrow and Chris Martin announced their separation on 25 March 2014. Certainly, social media also helps to spread these linguistic cults and trends – within a matter of hours, they can have a worldwide following. But the progression of change is much more exciting and revealing than the introduction of a few new expressions for reasons of fashion. One of the most exciting areas in historical linguistics now involves the research that looks at the social use of language and its role in language change. This is where we now turn our attention.

All changes occur in a social context

Language is a social product, yet curiously we had to wait until the 1960s before we had any real understanding about the social factors promoting language change. True, there were some famous nineteenth-century linguists who were curious about the speaking individual and society as a whole, but in the early twentieth century attention suddenly shifted away from speakers and fixated on the language system. The reason was the influential Swiss linguist Ferdinand de Saussure (1857–1913). Among his many contributions to the scientific investigation of language was the strict separation of **synchronic** (= static) and **diachronic** (= historical) perspectives. (We introduced you to these concepts in Chapter 1.) He was also a key architect of linguistic **structuralism**, a new way of thinking about language as a highly organised system of interlinking elements. Not only did this view privilege the underlying structure of language (Saussure's *langue*) above the actual use of language (Saussure's *parole*), it emphasised synchrony as the only legitimate approach to studying language. Saussure's way of looking at language was immensely successful, but it neglected the very facets of language that provide the clues to understanding change – namely, the social, behavioural and even historical aspects.

For Saussure and his followers, there was an immensely puzzling aspect to language: if language was this structured system of sounds, words and grammatical forms, how could it change without completely disrupting the system? Saussure himself painted a picture of a chess game to illustrate the conundrum. How could players continue to play the game if the rules were constantly changing? In the same way, how could language continue to be an effective vehicle of communication if it was ceaselessly on the move? This is especially puzzling when you consider the scale of some changes. Old English evolved into a completely different-looking Modern English, yet speakers continued to communicate effectively, even while these colossal changes were taking place. This puzzle is now referred to as the 'Saussurean Paradox'.

In fact, there is no mystery if language use and all its various forms are simply factored in as part of the structure of a language. It is through variation that change takes place – socially significant variant forms compete for attention. With time, speakers simply shift their allegiances and change follows. But, linguistic forms, clearly, first have to acquire some type of prestige before changes happen. These can involve a conscious shift towards an overtly prestigious form. As you saw in Chapter 10, overt prestige is based on norms of behaviour sanctioned and encouraged by mainstream society, and this is described as ‘change from above’. Not all change is inspired by overt prestige, however. Non-standard varieties inspire changes too, and this involves forms with covert prestige (or street cred) – this is ‘change from below’.

Here is a handy catchphrase to remember: variation is the synchronic (or now) aspect of change; change is the diachronic (or historical) aspect of variation. Like many catchy learning devices, it is a little misleading in its simplicity, in that it does suggest that all variation is indicative of changes underway. But this is not the case. Variation in speech patterns can be just that – variation. A good illustration is whether or not speakers pronounce a velar nasal [ŋ] or an alveolar nasal [n] at the end of words like *swimming*. So-called ‘G-dropping’, indicated by spellings such as *swimmin’* and *bringin’*, has fallen in and out of favour for a long time, and it continues to give off clear social signals.

A tug of love over many centuries

In early Modern English, ‘G-less’ pronunciations like *swimmin’* were widespread and considered quite acceptable. Cooper, an early orthoepist (or scholar of pronunciation), stated in 1685 that *coughing* rhymed with *coffin* (= ‘coffin’), and *coming* with *cummin*. Among Jonathan Swift’s rhyming couplets were pairs such as *fitting* and *spit in* (and the Dean of St Patrick’s wouldn’t have dropped his <g>s if it were a bad thing!). Also significant are spellings of *napkin* and *cushion* as *napking* and *cushing*. (Shakespeare was guilty of these!) They show us that pronunciations like *swimmin’* were entrenched. The now favoured [ŋ] pronunciation for *swimming* [swɪmɪŋ] arose through pressure of spelling and the clout of the written word – spelling indicated a velar nasal, so G-dropping fell from grace. (Of course speakers aren’t really dropping a [ɡ] here but changing the place of articulation of the final nasal, which spellings represent as <ng>.)

This was complicated variation, as shown by Benjamin Suggitt Nayler’s recommendations for early Australian English. His 1869 book, with the extravagant title (typical of the time) *Commonsense Observations on the Existence of Rules (not yet reduced to System, in any Work extant) regarding the English Language*, is full of remarks on pronunciation and references to ‘conspicuous blemishes in the pronunciation of thousands, and tens of thousands, who speak the English language’ and was clearly intended to help the ‘Schoolmaster (or Mistress) in the Colony’. On the matter of G-dropping, Nayler is clear (though complicated!): ‘What I wish my readers clearly to understand is this – that *not* to give “the ringing sound” at all, is *a defect*; while, to give it on two consecutive syllables, is *a trespass*; it should neither be *overdone* nor *underdone*.’ So Nayler is saying that to G-drop in words like *swimming* is a defect; but he recommends the [in] variant where the verb stem already ends in [n], and he indicates this in spellings such as *bringin’* and *singin’*. It is telling that he is less condemning of the ‘underdone’ pronunciation than one he describes as ‘overdone’ (where the second *ing* ‘is harshly rung’). This he labels a ‘tautophony … contrary to Good-usage – not to say *decidedly vulgar’*.

As we saw in Chapter 10, the American linguist William Labov was the first to show that the relationship between society, linguistic variation and change could be studied in a systematic way. His 1966 account of New York English explored a number of sociolinguistic variables that were used within the New York speech community, such as post-vocalic /r/ in words such as *beer* and *pork*, the pronunciation of <th> in words such as *thirty* and *this*, and the pronunciation of <ing> in words such as *cooking* and *eating*, as just described. His methods revealed that language change always occurs in the presence of socially relevant variation (according to parameters such as social class, gender, age and the contextual feature of style) – and it is directly observable. This new way of looking at language, variation and change sparked a whole new line of research into linguistic evolution.

For example, in the early 1990s, New Zealand linguist Janet Holmes investigated the incidence of T-flapping in the pronunciations of words like *atom* as [æɾəm] (remember that this is an example of assimilation – the /t/ is voiced and ‘flapped’ against the alveolar ridge, taking on the qualities of the surrounding vowels). Holmes showed that younger New Zealand English speakers flapped more than middle-aged speakers, the working class used it more than the middle class and men used it more than women. It was also more frequent in casual contexts (conversation) than formal contexts (interviews) – in other words, in contexts where people were **least** aware of their usage. In short, flapping was most established in the conversational style of young working-class males (who scored 82 per cent), and least established in the interview style of middle-class, middle-aged women (who scored 35 per cent). But, was this variation a sign of change and, if so, what sectors of New Zealand society were leading the change?

When change is underway, you expect younger speakers to use the feature more than older speakers, but the really telling clue here was that flapping occurred more frequently in the conversational speech of young middle-class women than young middle-class males. These women were not far behind the young working-class males, scoring a surprisingly high 70 per cent (against the 59 per cent of their young middle-class male counterparts). So, they appeared to be acting as a kind of spearhead for this vernacular change to infiltrate the middle class. In short, flapping was entering New Zealand English via working-class vernacular speech, but was spreading into middle-class speech via young middle-class women. More recent accounts of this variation show the feature to be now firmly entrenched in the speech of New Zealanders (and Australians) – flapping appears to be here to stay. This represents a move away from the more overtly prestigious pronunciation of /t/ towards its vernacular or 'non-standard' pronunciation.

Changes proceed gradually through the system

Change is not only socially structured; it is also linguistically structured. In sound change, for example, shifts are not purely mechanical processes that involve all relevant vocabulary items at once; rather, they are gradual, slowly spreading through the vocabulary and affecting different groups of words at different times. This is called **lexical diffusion**. All sorts of factors are at work to influence which words are changed and when. In the case of New Zealand flapping in the 1990s, for instance, word position, stress and even vowel length influenced whether or not a speaker was more or less likely to flap a /t/. A favourite environment was word-final as in *get eggs*. Phrases with /t/ in this position were like forwards in the front line of diffusion. Frequency was also a factor – common expressions like *sort of* were very susceptible to flapping (and we explain why in a moment).

This 'sneaky diffusion' of sound change through the lexicon produces the typical S-curve shown in Figure 11.6.

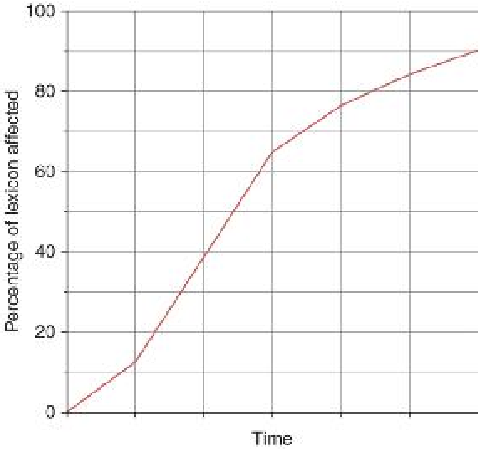


Figure 11.6 S-curve for linguistic change

A handful of words are affected at first (usually the frequent ones); the change then gathers momentum, but may later run out of steam. So, changes may not reach 100 per cent. Some sound changes simply peter out, and sometimes they are thwarted by a competing sound change (for instance, the conservative pronunciation of *great*, which never shifted to [i:] like *meat*).

Note that grammatical changes are subject to the same gradual dispersal as phonological changes. Take the case of the disappearing adjective endings. Phrases with *more* and *most* (*more tasty* and *most tasty*) are gradually pushing out forms with *-er* and *-est* (*tastier* and *tastiest*). Rather than a clean takeover, however, we find the usual variability associated with language change – but it is not random. Typically, adjectives of one syllable retain the endings (*bigger/biggest*). Adjectives of three or more syllables have lost the fight completely and require *more* and *most* (**beautifuller* and **beautifullest* sound awful). Adjectives of two syllables are in flux – some take the endings, some don't, some go both ways and not all speakers agree. Do you prefer *tastier* and *tastiest*, or *more tasty* and *most tasty*? Certainly, Shakespeare's *horrider* and *certainer* sound odd to us now, but how about *wickeder* and *wickedest* versus *more* and *most wicked*? These forms are the ones to watch.

11.4 The 'why' of change

The explanation for language change is one of the most interesting areas of historical linguistics, and in this chapter we have already touched on some of the driving forces. As you might imagine, they are numerous and complicated, and typically involve a network of intersecting pressures: psychological (the mental make-up of speakers), physiological (the production of language), systemic (the linguistic system with interacting components) and social (the speech community and the individual). Thrown into the mix are external or foreign influences, and we go on to address these in Chapter 12. There are even human wildcard factors to consider: the cultural preoccupations of speakers can be powerful triggers for dramatic and often unexpected changes. We are certainly not suggesting here that change is random, but it is hugely complex, with contingent internal and external factors (systemic and speaker-oriented), snowballing effects, and elements of chance. Language is a bit like climate and weather – not governed by deterministic law-like processes, but not frenzied or anarchic either.

We can't do justice to this enormous topic here, but we will illustrate just three contributing factors: physiology, society and frequency.

Physiological factors: Simplicity

> Language does what it has to for efficiency and gets away with what it can.
>
> Joseph Desmond O'Connor (1973), *Phonetics*

There have been all sorts of wacky wrong-headed theories put forward over the years for why the sounds of a language change, and in the case of the Australian accent speculations have been particularly wacky and wrong-headed. Some examples include a free-wheeling and adventurous spirit, a harsh climate, an outlaw heritage, a national nose inflammation through pollen or hay, a fear of opening the mouth on account of dust and flies, even ill-fitting dentures. Then, of course, there are those accounts that look to carelessness, sloppiness, laziness – and, according to one explanation, a kind of gross national inferiority complex. There is even one theory that attributes the Aussie accent to alcohol consumption. We have said that any change involves a complex of different contributing factors, and certainly speakers can use linguistic differences to highlight their ethnic character, but let us emphasise here that sound change is not caused by ethnic character, anatomy, climate, geography, bad dentistry – or alcohol.

One popular theory for which there is support is simplicity. You have seen how all languages show sounds dropping out, merging or assimilating over time. Such changes have been described as economy of effort, following the line of least resistance, human laziness and so on – but the end result is often a more efficient, more streamlined production. So, yes, many changes do result in simplification, but it's not quite so straightforward on the way through. English speakers stopped pronouncing the [k] in *knee* presumably because it was easier to say it this way – so why didn't German speakers do the same in *Knie* 'knee'? Simplicity also can't account for the fact that many changes result in greater complexity – for example, English consonants [ð] and [θ] (rather rare sounds) evolved from simple stops. Moreover, a change might simplify one part of the language but introduce complications elsewhere. You saw earlier how many of the irregularities in Modern English grammar (such as *was–were, foot–feet* and *mouse–mice*) were the fallout of sound changes that simplified English phonology at the cost of complicating its grammar. In short, when sounds drop out or assimilate, this can contribute to greater complexity overall. Simplicity is a factor, but it can't be the only one.

Social factors: The speech community

> Humans are more like stars than sponge cakes, since they group themselves into loose-knit clusters. Study of these social networks can reveal the intricate interlacing of human contacts. Potentially, they can show who influences who.
>
> Jean Aitchison (2013), *Language Change: Progress or Decay*

Early linguists noted that periods of dramatic linguistic change appeared to go hand in hand with social upheaval (epidemics, civil wars, revolutions, economic crises). The Black Death coincided with a number of major changes that took place in English during the Middle Ages. But why should this be? We now have a plausible explanation for this link and it is to be found in the social lives of speakers, in particular their membership of **social networks**, a concept we introduced in Chapter 10.

Just as Jean Aitchison describes in the quotation above, speech communities aren't like simple layer cakes, with the upper-class, middle-class and working-class layers piled up on top of each other. They are more like clusters of stars. As we noted in Chapter 10, networks vary in their density. In their day-to-day living, people interact with others in a number of different roles. However, in a dense network, someone might have a relative who is also a neighbour and the two might share the same place of work. This sort of repeated interaction across multiple roles with a small group of others is dense because of the high number of links between individuals.

Dense networks are great norm enforcers – quite simply, members of tight-knit networks are not going to want to speak differently from their mates. Such networks are most likely to be found in small and traditional communities with a high degree of social cohesion and, not surprisingly, they correlate with conservative speech patterns and lack of innovation.

Where there is little social cohesion between members of the speech community, social networks are said to be loose. These are more typical of people living in larger communities who have many external contacts. There is much less pressure in these contexts to be like one's neighbours, so these networks are more open to innovation. Changes are spread via the weak ties between different social networks – people (even those closely associated with one social group) can have ties with other groups and contact with these groups can affect their language. These speakers are the conduits for change.

So, back to why the bubonic plague should coincide with linguistic change. This was an epidemic that brought death on a scale comparable with a nuclear holocaust and it was made all the worse by a number of subsequent epidemics. Europe's population in 1430 is reputed to have been 50–75 per cent lower than it had been a century and a half earlier. Such devastation triggered mass emotional disorders and would have torn even the densest of social networks apart. It is precisely at such tumultuous times that variants are able to take off, spread and eventually embed themselves as long-term changes in the language system. It is as if changes are always lurking in the wings, but it is times of mayhem that enable them to catch on and flourish.

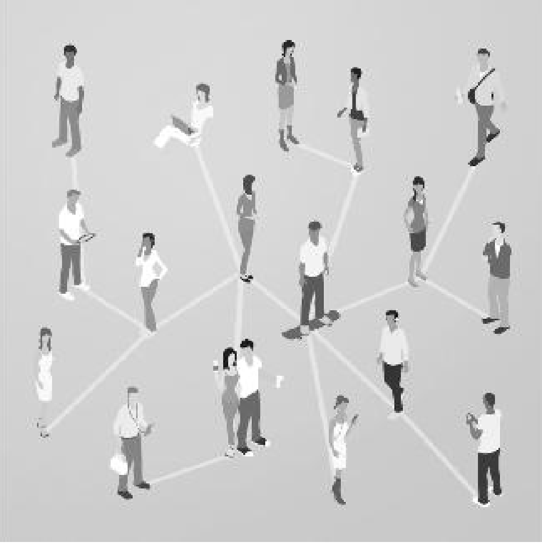


Figure 11.7 Social networks (the lines show the links between the people)

To conclude this section, you might contemplate how the internet, in particular social media, is now changing language. On the internet, people converse with each other in virtual communities just as they do face to face – except, of course, they are usually doing it in writing. The writing conveys information, but people are also always working on communicating their identities and group affiliations. As we described in Chapter 10, for members of virtual speech communities to express their identity (both personal and shared with other members of the group), they have to break away from the written standard and do their own thing. This is a great scenario for variation and change.

Frequency factors: The effects of repetition

Linguists now realise that an important motivating factor for language change is the frequency with which we use words and phrases. Many different aspects of our linguistic behaviour are shaped by repetition. Frequency leads to the reduction of meaning. Swear words and abuse terms lose their potency and wounding capacity. *Bloody* was once considered such an appalling word that it was rendered in print with asterisks (*b****y*). Now it is little more than a variant of *very*. This is a familiar story in vocabulary change: constant tugs of war between creativity and expressiveness on the one side, and a tendency for words to become routine and mundane on the other.

You have also seen how repetition has a profound effect on pronunciation. Individual words and strings of words that are common typically shorten over time – *Australia* becomes 'straya' [stræɪja]. Well-aged compound words reduce – *daisy* began life as 'day's eye'. Reduction is most obvious in formulaic expressions such as greetings and leave-takings – the ritualised string *God be with you* is ground down to *bye*. Pairs of words used frequently are also more likely to reduce; in normal rapid speech, *don't you* is pronounced as 'donchew' [dəʉntʃu]. This sort of repetition also plays an important role in the building of new grammar. Grammars will always reflect the linguistic habits of their speakers' daily discourse. High-frequency *going to* becomes the future marker *gonna* – it will (is gonna?) replace *will*, just as *will* replaced *shall*.

Aspects of lexical diffusion are the fallout of frequency effects. As described earlier, sound change is not a purely mechanical process that bulldozes its way through the lexicon, changing at once all eligible words in its path. It is gradual, slowly sneaking through vocabulary and affecting different groups of words at different times. Common usage words are typically the first to be affected. English is currently losing the vowel before the [ɹ] in endings such as *-ery*, *-ary* and *-ory*; words such as *every* [evɹi] and *scenery* [sinɹi] are at the forefront of this change, and less common words like *mandatory* are slower to show the change.

There are a number of reasons why repetition leads to the reduction of form. One is that frequent or everyday words will often crop up in casual settings where short-cuts and reductions are tolerated more – a lot becomes predictable and so can be taken for granted when you're chatting with someone you know well in a familiar environment. Words of high frequency are generally easier to understand, so listeners won't require the same amount of phonetic detail. Reduction is also the result of the automation that follows from the repetition of articulatory movements. Linguists have shown that when gestures of articulation become automatic, this triggers a reduction in the timing and magnitude of these gestures. The speed of execution automatically increases and the transition between sounds becomes more fluent – movements might even overlap. A sailor who says the word *captain* often will reduce it to 'capm' [kæʔm]. We see this sort of thing in many repetitive activities, not just linguistic gestures. Think of the salute of that same sailor – there will be short-cuts in that gesture too.

In grammatical change, everyday words are again exceptional, but for the opposite reason – they are typically the **last** to change (and may never change). Verbs provide a good example. We described earlier how English verbs are gradually adopting the regular *-ed* ending to indicate past time. Those that have resisted are the high-frequency verbs like *go–went*. As linguist Joan Bybee describes, irregular patterns like *go–went* have to be stored individually in the memory. Should memory fail, the regular rule steps in and (via analogy) we get *goed*, and we can readily understand this form, because we are dealing with the default rule. However, *goed* doesn't triumph (as *climbed* did over *clomb*) because its high frequency strengthens it against the (analogical) changes occurring elsewhere. Such irregular verb forms are listed individually, and every time we access them they get more entrenched and able to resist the regularising forces outside. Other irregular verbs, such as *heave–hove, smite–smote* and *beget–begat*, are simply not frequent enough to have stuck in our brains. They've lost out.

So, we are left with an apparent paradox. Common everyday words are **first** to be affected in sound change but are **last** to be affected in grammatical change. Nothing in language is ever straightforward!

11.5 Language families and establishing a genetic relationship

> There is no tracing the connection of ancient nations, but by language; and therefore I am always sorry when any language is lost, because languages are the pedigree of nations.
>
> Samuel Johnson, 18 September 1773, in James Boswell, *The Journal of a Tour to the Hebrides*

The previous sections have focused on linguistic change. We now turn our attention to the question of **genetic affinity** and how we might uncover the linguistic relationships that arise as a consequence of change. To set the scene, we present the opening lines of the Lord's Prayer at different stages of English:

- Old English (c. 1000): *Fæder ure þu þe eart on heofunum, si þin nama gehalgod.*
- Middle English (c. 1400): *Oure fadir that art in heuenes halowid be thi name.*
- Early Modern English (c. 1611): *Our father which art in heaven, hallowed be thy Name*.

Now compare the English and German versions:

- Old High German (c. 830): *Fater unser thu thar bist in himile, si hiheilagot thin namo*.
- Modern German: *Unser Vater in dem Himmel. Dein Name werde geheiligt.*

There are striking similarities between English and German in lexicon, sounds, syntax and morphology – which is not surprising, given that they are close relatives. The similarities are because they have evolved from a single parent language: they are daughter languages of Proto-Germanic. The differences that exist between them are due to language change, so the further back we go in time and undo the changes, the more striking the similarity becomes. In fact, speakers of Old High German and Old English would have been able to understand each other.

Now consider the Welsh and the Hungarian versions of the Lord's Prayer:

- Welsh: *Ein tad, yr hwn wyt yn y nefoedd: sancteiddier dy enw.*
- Modern Hungarian: *Mi Atyánk, ki vagy a mennyekben, szenteltessék meg a te neven.*

Glancing at the Welsh and Hungarian versions, there is nothing to suggest that either language is genetically related to German and English. We know that Welsh is a relative, but Hungarian is not. How can this be, when on the surface they look just as different? How can we justify the claim that certain languages are related while others are not?

How to confirm genetic relationships

It is not so much similarities that confirm the relatedness of languages (in the case of Welsh, these aren't all that apparent), but the existence of systematic differences. Consider the Dutch, Swedish, English and German words in Table 11.11. Here is a set of words that have the same meaning and are also phonetically close. Most importantly, the differences between them are systematic. Where Dutch, German and Swedish have *d*, English has *th*. This is known as a **sound correspondence**.

Table 11.11 Cognates in Germanic languages

[illegible]
[illegible]
[illegible]

These kinds of words are called **cognates** – words similar in both form and meaning because they are historically (or etymologically) related. They are not the result of borrowing because we can find hundreds of such cognate sets, and borrowings don't normally occur on such a large and systematic scale. Furthermore, the words are in everyday use. Why would speakers borrow 'brother' or 'you'? Languages usually borrow if they need terms to express new concepts. It is not likely that the similarities are due to chance, either. Given that individual words are arbitrary, there is little possibility that speakers of these languages would independently arrive at the same or similar sound patterns to represent the concept. We have no choice but to assume that these words are connected in origin.

Using a technique known as the **comparative method**, we can even reconstruct the Proto-Germanic words in Table 11.11. The asterisk here means they are hypothetical or unattested forms (see Figure 11.8). These ancestral words represent a relationship between sound and meaning that is carried through to each of the modern Germanic languages, despite at times drastic sound changes.

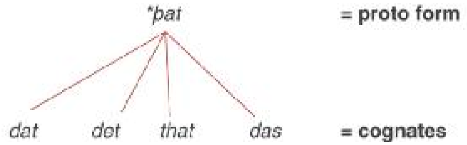


Figure 11.8 Ancestral ‘that’ and its daughters

If we are serious about reconstructing the Germanic parent, we would seek out forms from the earliest available evidence of these languages. The data set in Table 11.12 gives cognates from Gothic (the earliest attested Germanic language), Old Norse (early Scandinavian) and Old English. Consider only the vowel sounds.

Table 11.12 Cognates from early Germanic

Gothic	**Old Norse**	**Old English**	
stains	*steinn*	*stan*	‘stone’
hails	*heill*	*hal*	‘hale’
haims	*heimr*	*ham*	‘home’
gaits	*geit*	*gat*	‘goat’
hlaifs	*hleifr*	*hlaf*	‘loaf’
[ai]	[ei]	[a]	

The similarity between the languages is obvious and, again, too striking to be the result of chance. But, most importantly, the differences between the languages are systematic. Gothic [ai] corresponds to Old Norse [ei] and Old English [a]. The correspondence is regular and can only be accounted for if we assume that it results from a time of mutual development – a common ancestor that developed one way in Gothic, another way in Scandinavian and another way in English.

One final point to emphasise here is that sound change is regular. The original sound [*ai] changed to [a] in Old English, and it did so in every word under the same condition. Furthermore, in moving forward in time from Old English, all these words later show the same vowel shift to [əʉ] in Australian English and [ɐʉ] in New Zealand English. The regularity of sound change is a cornerstone of this method; it is what allows us to compare languages and to reconstruct ancestral forms. Without sound change regularity, the comparative method would not be viable.

How do language families arise?

The concept of genetic relatedness between languages had been bandied about for a long time, but it wasn’t until a lecture in 1786 that the idea really took off. Sir William Jones argued that the ancient languages Sanskrit, Ancient Greek and Latin were related, and could be traced back to ‘some common source, which, perhaps, no longer exists’. Jones was arguing for **Proto-Indo-European** (known fondly as PIE), the parent of most of the languages of Europe, south-western Asia and northern India (assumed to be spoken sometime around 3000 bce). It is a bit unfair to give all the credit to Jones here, since others had made similar observations. But, it was Jones who really advanced the idea of a hypothetical proto-language – up until then, most scholars had been arguing for one of the existing classical languages as the parent. PIE has left us no written records because writing hadn’t been invented then, but we know it must have existed because of the regular correspondences between the sound systems of these languages.

Jones’ ideas were enthusiastically taken up by other scholars, and before long the systematic study of language evolution was undertaken for the first time. It was this work that led to the classification of all the Indo-European languages into the sort of genealogical tree you see in Figure 11.9.

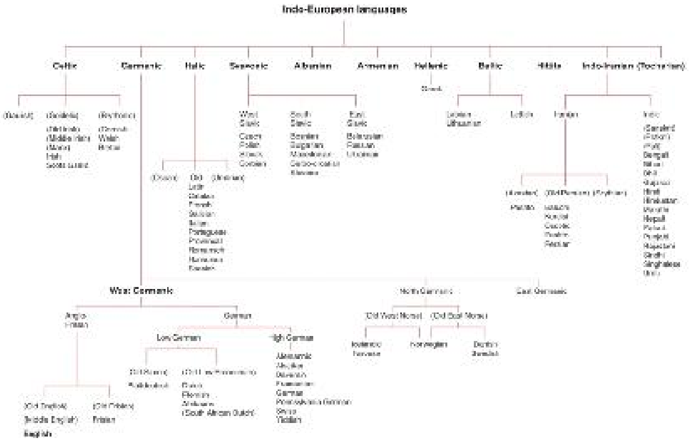


Figure 11.9 Indo-European languages

To get an idea of how language families evolve, imagine the conditions that might have given rise to the Germanic family of languages. Roughly around 1500 bce, a bunch of speakers left the Indo-European homeland – perhaps there were disputes or the population got too big for the area to support, or maybe they simply had an adventurous spirit. All languages change, but they are more likely to change when they are in different places. In this case, the Germanic group developed a number of distinctive characteristics, which then set it apart from the other Indo-European groups. The most significant of these Germanic innovations was something we now know as the **First Germanic Consonant Shift** (or **Grimm's Law**, after one of the famous brothers Grimm, who discussed it in his 1822 grammar of German). Remember that these consonant shifts occurred in ancient times before the Germanic language split up into all the different branches. They are what served to make the Germanic group distinct from the other Indo-European languages.

Proto-Indo-European	**Germanic**
(voiceless stops) p, t, k	(voiceless fricatives) f, θ, x (h)
(voiced stops) b, d, g	(voiceless stops) p, t, k
(breathy voiced stops) bh, dh, gh	(voiced stops) b, d, g

To illustrate these innovations, in Table 11.13 we present some cognates from Latin, Sanskrit and Old English. The Latin and Sanskrit words represent the original sounds; the Old English ones represent the Germanic innovations. The sound correspondences are given each time in bold and the ancestral phonemes from which they derive are given in the first column.

Table 11.13 Indo-European cognates

PIE	**Latin**	**Old English**
*p	*piscis* 'fish'	*fisc*
*t	*tres*	*threo*
*k	*cord* 'heart'	*heorte*
*b	*bursa* 'purse'	*pursa*
*d	*dent-* 'tooth'	*toth*
*g	*genu* 'knee'	*cneo*

PIE	**Sanskrit**	**Old English**
*bh	*bhárāmi* 'I carry'	*beran* 'bear'
*dh	*bandhanam* 'bond'	*bindan* 'bind'
*gh	*ghnánti* 'he strikes'	*gūþ* 'battle'

There were many more sound changes on the road from PIE to Germanic than just the ones represented here. There were many grammatical changes too – for example, Germanic developed its own class of verbs indicating the past with some sort of dental suffix (such as English *-ed* and German *-t*). With time, there were more and more changes until Germanic eventually parted company from Indo-European. Through the combination of time, geographical separation and the processes of change, a single language can split into dialects and, over a long enough period of time, these dialects can turn into distinct languages.

And so Germanic split off from Indo-European. The number of speakers increased and, as the language spread geographically, it changed in different ways in the different parts of Europe, giving the three main branches you can see on the tree: North, East and West Germanic. North Germanic is the parent of the modern Scandinavian languages (except Finnish, which isn't Indo-European). East Germanic is the parent of Gothic and its descendant Crimean Gothic (both long dead). West Germanic underwent a number of additional splits, eventually sprouting the modern groupings you can see here.

What does PIE look like?

Unfortunately, speakers of PIE weren't big writers – or at least, they left us no written texts. So, what was it like? The German linguist August Schleicher (1821–68) was the first to attempt a systematic reconstruction of the sounds and forms of this proto language; to make it even more real, he went to the lengths of recreating an actual text in this language. Called 'The Sheep and the Horses', the short parable tells the story of a shorn sheep that encounters a group of unpleasant horses. Here is Schleicher's parable with a Modern English translation.

Avis akvāsas ka

Avis, jasmin varnā na ā ast, dadarka akvams, tam, vāgham garum vaghantam, tam, bhāram magham, tam, manum āku bharantam. Avis akvabhjams ā vavakat: kard aghnutai mai vidanti manum akvams agantam. Akvāsas ā vavakant: krudhi avai, kard aghnutai vividvant-svas: manus patis varnām avisāms karnauti svabhjam gharmam vastram avibhjams ka varnā na asti. Tat kukruvants avis agram ā bhugat.

The Sheep and the Horses

A sheep that had no wool saw horses, one of them pulling a heavy wagon, one carrying a big load and one carrying a man quickly. The sheep said to the horses: 'My heart pains me, seeing a man driving horses.' The horses said: 'Listen, sheep, our hearts pain us when we see this: a man, the master, makes the wool of the sheep into a warm garment for himself. And the sheep has no wool.' Having heard this, the sheep fled into the plain.

Schleicher was the first linguist to use the family tree as a way of graphically representing historical changes involving gradual divergence over time and eventually resulting in the formation of new languages. Although his creations of PIE were condemned as flights of fancy, reconstructions such as 'The Sheep and the Horses' offered a compelling demonstration of the comparative method at work.

Mapping change as trees and waves

Even as early as the nineteenth century, many scholars were unconvinced by the family tree as a way of describing the development of languages. Its 'single-parent source for every language' approach not only failed to accommodate the fuzziness of language and dialect boundaries, but also failed to capture the fact that languages (related or otherwise) may continue to influence each other, sometimes over long periods (we will see this clearly in Chapter 12).

Johannes Schmidt (1843–1901) proposed what has now come to be known as the 'wave model' or 'wave theory' (*Wellentheorie*). He claimed, correctly, that linguistic changes were restricted spatially and his wave metaphor captured the fact that new features of a language can spread from a central point in continuously weakening concentric circles – much like the waves created when something is thrown into a body of water. By showing that each sound change had its own territory, Schmidt introduced what has become the modern concept of the **isogloss**, the geographical lines indicating the domain of sound changes.

These isoglosses are drawn on a map to indicate where a usage changes – or how far the ripples from the stone extend. Consider the following map showing the boundaries between Low German ***maken*** and ***appel*** and High German ***machen*** 'make' and ***Apfel*** 'apple' (changes caused by something called the Second Germanic Consonant Shift). When lots of isoglosses like these appear in the same place, we recognise a dialect boundary. For example, dialects north of the lighter line in Figure 11.10 have the original [k] and those to the south have the newer [x] in words like *maken/machen*. Dialect maps like the one in Figure 11.10 allow linguists to reconstruct likely population movements and contact between language varieties in the region.

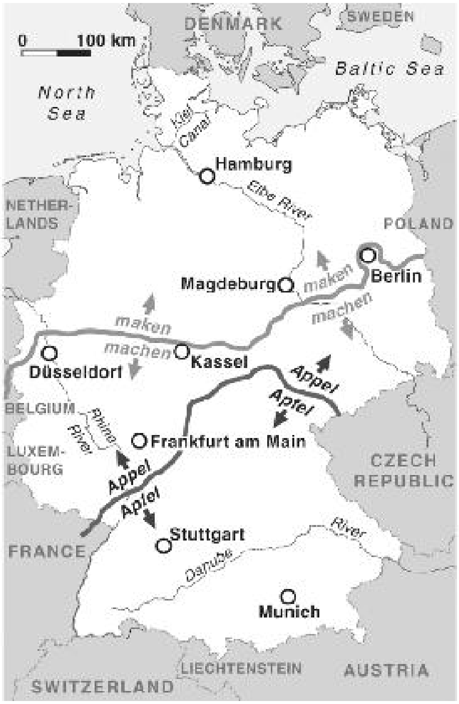


Figure 11.10 Isoglosses of the West Germanic dialects

In fact, the 'family tree' and 'wave' models are compatible with each other; Schmidt himself saw his model as supplementing the standard family tree, simply by providing a more complicated version of the single splits offered by family trees.

11.6 Attitudes to change

> What is a weed? A plant whose virtues have not yet been discovered.
>
> Ralph Waldo Emerson (1878), *Fortune of the Republic*

For most speakers, change is perfectly fine (it might even be interesting), as long as it takes place well and truly in the past. Most people appear to be fascinated by the origins of words and the stories that lie behind the structures in their language, but they experience a real squeamishness about any changes happening under their noses. Typically, linguistic innovation is seen as change for the worse – an indication that the language is going to the dogs.

Take a straightforward example. In March 2014, the *Oxford English Dictionary* published its new list of words, and among them was the verb *to toilet-paper* 'cover (a building, trees, and so on) with toilet paper'. This word reflects that handy ability of English to convert words into other parts of speech without adding anything in the way of an affix. This is not a new trend by any means, but still a convenience not all speakers fully appreciate, judging by the negative reactions to this neologism (and also to news reports of those sporting heroes who *podiumed* and *medalled* during the 2014 Winter Olympics). New conversions typically provoke hostility simply because they are unfamiliar. But, with time, recent newcomers such as *to toilet paper, to podium* and *to medal* (if they survive) will sound as everyday as venerable oldies such as *to contact* and *to hand.* And the next generation of English speakers will be puzzling over what possible objections there could have been to these words. Of course, they will have their own linguistic battles to wage.

Why do people feel antagonistic towards language change? Partly it's their nostalgia for things of the past – the feeling that what is on offer today is not quite up to scratch. So, their natural reaction is to condemn change. Many linguistic pet hates also have deep social underpinnings. The verb *to contact* (from the noun *contact*) was once jargon, and 80 years ago there was ferocious resistance. An outraged public used descriptions like 'lubricious barbarism' for this linguistic newcomer. In the same way, today's neologisms *to beverage* and *to impact* also seem a little self-important. So when people grumble about these expressions, they are often not rejecting them outright. It is more that they don't approve of the kind of people who use this sort of language, a topic to which we return in Chapter 15.

The metaphor of the weed is an appropriate one to draw on here. Definitions of weeds are many: 'a plant growing where we do not want it'; 'a plant whose virtues are yet to be discovered'; 'a plant growing out of place'; 'a plant that you do not want'; 'a plant you hate' and so on. Clearly, weeds are context-specific; it depends entirely on location and on time whether something is classified as a weed. And so it is with what some might describe as the weedy aspects of language. Linguistic weeds are structural features of the language whose virtues have yet to be realised. They are the pronunciations we don't want, the constructions that are out of place, the words we create but hate. Like weedy plants, they are entirely location and time specific. One speaker's noxious weed can be another's garden ornamental. Whether they are in gardens or in languages, weeds are totally centred around human value judgements.

What many see as slipshod pronunciations, sloppy grammar, irritating coinages and new-fangled meanings are what provide the basis for real change. Some will drop by the wayside, it is true. Some will remain as variations, as we've just discussed. But, there will be others that catch on, are used more and more, and eventually become established. The burning question is always which ‘weedy’ tendencies are likely to survive and transform into tomorrow’s respected and rewarding species.

The All Blacks are versing the Wallabies

> All my primary-school-age grandchildren, from as far apart as Brighton and Berwick, are using ‘vers’ as a verb: ‘Did you vers him?’ ‘I am versing you.’ ‘Will you vers me?’ ‘Did they vers each other?’ … Will teachers be powerless to stop the spread of such usage? Do you think it will quickly become part of the language?
>
> Email received 3 March 2010

In Chapter 4, we included this new verb *to verse* ‘to compete against’ as an example of backformation. It is formed on the pattern of existing verbs like *to curse* and is another product of analogy, but working in the opposite direction (starting from the complex form, not the simple form).

curses : *curse* as *versus* : ?
? = *verse*

Here the [əz] of *versus* has been reinterpreted as the present tense ending *-es*, and we have a brand-new verb, *to verse*. Will it be part of the language? Well, it’s already made it into the *Macquarie Dictionary* – much to many people’s horror. The lexicon is the most unstable aspect of a language, and dictionary makers are constantly having to redraw the admission and exclusion boundaries for marginal vocabulary items. If they didn’t, people would stop using their dictionaries. Yet, when new entries like *to verse* appear, there are often howls about declining educational standards. It’s all a lovely illustration of our human capacity for doublethink.

Further reading

There are some wonderful resources on linguistic change; we recommend Aitchison (2013), Campbell (2004), Crowley and Bowern (2010), Hock and Joseph (1996), McColl (2007) and Bauer (1994) (which focuses on English). See Johnston and Schembri (2007) for a discussion of the phonological processes in Auslan, and Labrune (2012, 2014) for processes in Japanese. Grammaticalisation is now a hot topic, and there are some excellent accounts around: see Hopper and Traugott (2003), and papers in the journal *Language Sciences* (vol. 23, 2–3, edited by Campbell and Janda (2001), a special issue devoted entirely to this topic). Greenberg (1960) calculated the ratio of number of morphemes per word; see also Lyons (1968) and Haspelmath (2002). The study of New Zealand flapping is Holmes (1994) and more recently Silby (2008). Lesley and James Milroy pioneered the work on modelling language contact between individuals; for their ideas about social networks and change see L. Milroy (1987) and J. Milroy (1992, 1993). Joan Bybee (2003a, 2003b) has researched the many different aspects of our linguistic behaviour that are shaped by repetition.

Exercises and discussion points

1 Phonological change – English

Explain and name the phonological processes that bring about the pronunciation of the following words in English. Try to think of at least one other example that illustrates the same process.

(a) Italian *colonella* ‘little column’ as Modern English *colonel* [kɜ:ɹnəl] or [kɜ:nəl]

(b) Modern English *ask* from Old English *acsian*. (Many modern speakers now *aks* – why do you think this is?)

(c) Modern English *drench* from Old Germanic **drankjan* ‘to cause someone to drink’ (compare *drink* < **drinkan*)

(d) the word *glamour* from an originally Scottish pronunciation of *grammar* (< Old French *gramaire*)

(e) the colloquial pronunciation of *seven* as [sebm] and *eleven* as [əlebm]

(f) the colloquial pronunciation of *Can I go*? as [ŋægou] (there are many processes here)

(g) the pronunciation [probli] for *probably*

(h) the pronunciation [bjʉdi] for *beauty*

(i) pronunciations of *memory* as *membry* [membri] and *youngster* as *younkster* [jɐŋkstə]. (Your explanation should also explain the phonological features of the sounds that are inserted.)

(j) a speaker intending to say *discussing shortly* who says *discushing*

2 Phonological change – languages other than English

Explain and name the phonological processes that bring about the pronunciation of the following words. If you can, think of an English example that illustrates the same process.

(a) Modern French *chercher* [ʃerʃe] from Old French *cercher* [sertʃer] 'search, look for' (from which English borrowed *search*) instead of the expected form [serʃe]

(b) Spanish [paɣaðo] 'pacified' and dialectal [paɣao] from Latin *pacatum* [pakatum]. (Only consider consonants here)

(c) Icelandic *finna* 'find' (< **findan*); *gull* 'gold' (< **gulþ*); *munn* 'mouth' (< **munþ*); *unna* 'love' (< **unþan*)

(d) Latin *flores* < pre-Latin **floses* (compare English *floral*)

(e) Old Irish [ma:l] 'prince' < Pre-Old Irish *[magl]; Old Irish [kene:l] 'gender' < Pre-Old Irish *[kenetl]

(f) Modern German word *Kartoffel* 'potato' from *Tartuffeln* (originally Italian *tartuffeli*, named after the truffle)

(g) Turkish *iyidir* 'it is good'; *kizdir* 'it is a girl'; *süttur* 'it is milk'; *kuştur* 'it is a bird' (explain the consonant and vowel changes in the suffix)

(h) Australian Kriol *apinum* from English *afternoon* [ɐftənʉːn]

(i) Australian Kriol *silip* from English *sleep*

(j) in some non-standard dialects of Fijian, the pronunciation [tinana] 'his/her mother' as [tʃinana]

3 Grammatical change – English and other languages

(a) In some dialects of English, the pattern *bring/brought/brought* has become *bring/brang/brang* or *bring/brang/brung*. Name the process this change illustrates.

(b) Middle English had *help* 'present tense', *holp* 'past tense'; Modern English has *help, helped* for these. Name the process this change illustrates.

(c) The analogical spread of the plural *-s* has been a real success story – so successful, in fact, that even the *s* on the end of some noun stems has been 'wrongly' interpreted as a plural, and a new singular then 'backformed'. Identify the source of the backformations in the following: *gentry, asset, eave*.

(d) Briefly explain Talmy Givon's catchphrase 'Today's morphology is yesterday's syntax'. Support your discussion with examples.

(e) Explain why the following examples are exceptional in the way they form their plurals: *Maple Leafs* (= 'Canadian ice-hockey team'), *still lifes* (= 'paintings of inanimate objects'), *Sabertooths* (= 'a type of tiger').

(f) The following letter by A.P. Taskunas (Sandy Bay, Tasmania) appeared some years back in *The Australian Magazine*:

> Shame on you! You've killed the accusative ('Deaths in the family: Who killed who? – Sept 25–26), one of the most endangered species in our language. Whom can we enlist to protect it?

What is A.P. Taskunas protesting about here, and how would a linguist respond to the outburst?

(g) In Australian Kriol, *gotta* has a variety of functions. Two are exemplified in the following sentences:

> *If you kick him here you gotta kill'im*. 'If you kick it there, you will kill it.'
>
> *One big plate gotta three little candle*. 'A big plate with three little candles.'

Explain the grammatical functions of *gotta* in these examples and name the change that is illustrated here.

4 Discussion questions on grammatical change

(a) In 1749, Benjamin Martin wrote, 'No language … can ever be permanently the same, but will always be in a mutable and fluctuating state, and what was deem'd polite and elegant in one age, may be accounted uncouth and barbarous in another'. What did he mean by this?

(b) Some linguists argue that grammatical change is goal-driven – in other words, languages develop grammatical categories (such as future) because they need them. Outline one reason why this cannot be the case. Give an example to support your answer.

(c) Simplification is often given as a reason for sounds to change. However, there are problems with this as an explanation for change. Describe one of these. Provide a supporting example.

(d) Frequency of use has a profound effect on linguistic behaviour

– frequent everyday words are first to be affected in sound change but last to be affected in grammar change. With examples, explain this paradox.

(e) As we noted in section 11.2, in 1921 Edward Sapir wrote, ‘A linguist that insists on talking about the Latin type of morphology as though it were necessarily the high-water mark of linguistic development, is like the zoologist that sees in the organic world a huge conspiracy to evolve the race-horse or the Jersey cow.’ What did he mean by this?

5 Genetic relationship

Examine the sets of words (in phonetic form) below from four Germanic languages and answer the questions that follow.

English	Dutch	German	Swedish
[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]

(a) What is the name given to such sets of words?

(b) Provide the sound correspondences for the vowel sounds.

(c) Give the Dutch, German and Swedish words that would correspond to the English word *brown*.

(d) How do we account for the striking similarities between the vocabulary items here? Could they be due to chance or borrowing?

(e) Draw a mini-family tree to represent the relationship between these languages.

Research project

History of language

Choose a language and do extensive reading on the historical linguistics of that language. Write up your findings, with discussion of the following points:

- attested stages of the language – consider, for example, early inscriptions, texts and stages in the development of the orthography/ies

- major developments in the system of the language – for example, sound changes, morphological changes and syntactic changes
- connections with other languages – include genetic/historical relationships and contact/convergence relationships (type and direction).

For more activities, please go to www.cambridge.edu.au/academic/linguistics.

12

Languages and cultures in contact

◆

Locutions are like seedlings; transplanting makes them better and stronger.

Michel de Montaigne (1572–80), ‘On Some Lines of Virgil’, in *The Complete Essays* (trans. by M.A. Screech, 1987)

12.0 Introduction

There are all sorts of reasons for languages to come together. Colonisation brought English into contact with the Indigenous languages of Australia and New Zealand; mobility and migration have now introduced more than 40 different ethnic groups into both these countries. Business was a trigger for the formation of roughly 60 English-based pidgins and creoles along the trade routes of Asia, Africa, the Caribbean and the Pacific, and changes in political boundaries between countries or states are constantly breaking down language borders and creating new ones. When languages come together for whatever reasons, all sorts of interesting things can happen. There is even a separate branch of linguistics that studies the fallout of this interaction, called (not surprisingly) contact linguistics.

The focus of this chapter is the complex interaction of internal and external factors that come together to shape the linguistic outcomes of contact. These outcomes can be everything from the incorporation of occasional vocabulary items to the creation of brand-new hybrid languages; it might entail the happy coexistence of two (or more) languages, or the dominance of one language over others. The features of each language, including how different or similar the languages are (for example, it is easier for languages to incorporate constructions that are similar to ones they already have), can have a significant impact on what happens. Social and psychological factors, including the length and intensity of the contact, and the power and prestige of relationships, are also involved. Also relevant is that different levels of language structure (for example, phonology, prosody, grammar and lexicon) are differentially sensitive to these factors, and have a bearing on the nature and direction of the transfer of features, structures and rules from one language to another.

We also can't talk about language contact without considering the significant losses to linguistic diversity across the world over the last few decades, as a few majority languages have increasingly gained dominance. This diversity goes well beyond the structural aspects of language. It encompasses the world-view of the speakers of those languages remaining in the world – including what their grammars make them pay attention to, and the different ways they end up allocating and organising reality. There is a lot that disappears with the sounds, meanings and grammar of those vanishing languages.

12.1 Types of contact between languages

> Intercourse between different peoples has always carried new words from one place to another, just like articles of trade, after which they either take root or disappear, depending on whether they are admitted or rejected by usage.
>
> Baldesar Castiglione (1528), *The Book of the Courtier* (trans. by George Bull, Penguin, 1967), Dedication

When speakers of one language come into contact with speakers of another, some people might become fully bilingual as children, while others learn the second language as adults with varying degrees of success. Worldwide, the majority of people are either bilingual (competent in two languages) or multilingual (competent in more than two languages). To provide some idea of the extent of the multilingualism in some speech communities, consider one of our case-study languages, Turung. Turung is part of an incredibly complex mosaic of languages in north-east India. It is spoken in the district of Balipathar, which covers about three square kilometres and has a total population of only 1500 people. No fewer than thirteen languages are spoken there, representing five different language families. These are summarised in Table 12.1.

Table 12.1 Languages spoken in Balipathar

Language	**Number of speakers**	**Language family and subgroup**
Aiton	50	Tai-Kadai (Southwestern Tai)
Turung	20	Tibeto-Burman (Jingpho)
[illegible]	[illegible]	Tibeto-Burman (Bodic)

Garo	Tibeto-Burman (Bodo-Koch)	25
Bodo	Tibeto-Burman (Bodo-Koch)	1
Kachari	Tibeto-Burman (Bodo-Koch)	1
Kuki	Tibeto-Burman (Mizo-Kuki-Chin)	1
Mizo	Tibeto-Burman (Mizo-Kuki-Chin)	3
Munda	Austro-Asiatic (Munda)	450
Oraon	Dravidian (Northern)	100
Sadri	Indo-European (Indo-Arayan)	800
Assamese	Indo-European (Indo-Arayan)	All
Nepali	Indo-European (Indo-Arayan)	200

Three additional languages are used in the religious traditions followed in this area (Pali, Tibetan and Burmese), and there is also a place for each of the national languages (English and Hindi), along with the most dominant minority language in the larger area (Karbi, a member of the Tibeto-Burman Mizo-Kuki-Chin sub-group).

In striking contrast are the (less usual) monolingual speech communities of the world where only one language is used. In countries such as Australia, Britain, Japan, France and Germany, for example, children will usually become native speakers of only one language. Of course, there are usually quite large groups of people in these countries that do make regular use of other languages; in Australia, for example, there are pockets of quite considerable linguistic diversity. And although people may have little or no face-to-face contact with speakers of other languages, foreign features (usually words and expressions) can also be transferred by mass media and via language teaching. So, even overwhelmingly monolingual countries will be affected by outside influences.

The ‘monolingual mindset’

> Australia distinguishes itself from the rest of the world, which generally requires children to learn a second language … and at least gives them the opportunity to learn a third language.
>
> Michael Clyne (2008), ‘The Monolingual Mindset as an Impediment to the Development of Plurilingual Potential in Australia’

The English-speaking world is generally monolingual. Australia might be a multicultural and multilingual society, but most Australians are content to speak only English. And, while they cheerfully celebrate the country’s multiculturalism, they pay little or no attention to its multilingualism. Hence the popular joke among migrants to Australia:

> What do you call someone who speaks three languages?
> Answer: ‘Trilingual.’

> What do you call someone who speaks two languages?
> Answer: ‘Bilingual.’
> What do you call someone who speaks only one language?
> Answer: ‘Australian.’
>
> (This joke is told about various cultures and has many potential answers!)

Michael Clyne was a linguist who spent his life challenging what he called the ‘**monolingual mindset’** – the belief that monolingualism is the desired norm for countries like Australia and that educators should concentrate on the teaching of English and not waste valuable resources on other languages. As Clyne emphasised, it is a mindset that fails to recognise the powerful economic, political, social and cognitive motivations for multilingualism, and also neglects to capitalise on the linguistic diversity that exists in Australia and around the world.

The monolingual mindset is an example of a language ideology – a concept we explore further in Chapter 15. In Australia, the ideology of the monolingual mindset devalues languages other than English. This ideology has clear policy outcomes, including the failure of the education system to provide students with opportunities, incentives and rewards for studying certain second and third languages. Both the ideology and the policy have practical and psychological effects in the community, reducing language maintenance even in families where maintenance would have many benefits. They also limit opportunities for students from English-speaking backgrounds to develop cognitive and intercultural communication skills, and a better understanding of the world around them.

The monolingual mindset is not the only way things could go. In Aboriginal Australia, Papua New Guinea and, indeed, across most of the world, speaking two languages is considered to be the basic competence an adult needs in order to get by, so possessing abilities in at least three languages is quite normal. Consider Felicity Meakins’ description of language use among Gurindji people in Kalkaringi:

> Older people at Kalkaringi continue to speak a traditional language (Gurindji), however younger speakers speak a new youth variety (Gurindji Kriol) which systematically combines elements of Gurindji and Kriol. Some Warlpiri is also spoken; English is the main language of the school, clinic and other government facilities; and Kriol is spoken by Aboriginal visitors. In addition, people rapidly switch between these languages in various ways.

Direct contact is not required for linguistic features to cross borders. The fact that English is currently the largest (and most powerful) **lingua franca** (or common language) places it in an interesting position – and this, of course, is what encourages the complacent ‘monolingual mindset’. Because English has dominated areas such as science, diplomacy, education and business, it has made contributions over the years to the lexicons of many languages; some originally English words (such as *okay*) are now so widespread as to be truly international.

The outcomes of language contact

There are over 7000 languages spoken worldwide, but only in around 200 countries, which means that there is a lot of contact going on between languages (this figure is based on Ethnologue’s list published by SIL International). In recent times, much more contact has been occurring, and in new and increasingly complex forms. Never before have we seen so many people from so many different linguistic backgrounds wanting to interact. The result has been to foster new forms of language diversity in societies all over the world (‘super-diversity’, as it now called). Massive flows of people, including tourists, refugees and migrants, have produced an unprecedented intermixing of people and cultures. Languages and cultures at the local level have been changed irrevocably by this international movement of people. On top of this, we also have new technologies of communication now offering novel opportunities for interaction across the globe.

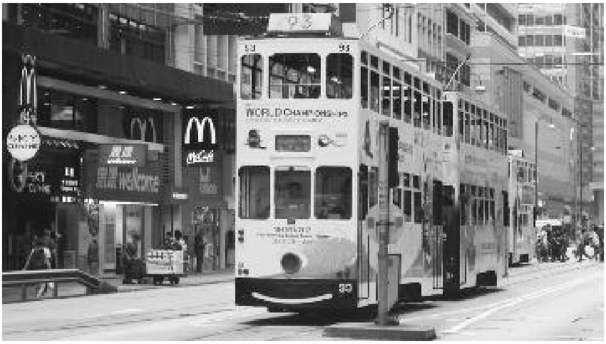


Figure 12.1 Languages in contact – a typical Hong Kong street

Metrolingualism

Major cities all over the world are key places in which languages meet and mix. Traditional approaches to language contact have focused on how languages are allocated to people in these settings, and the ways in which people within these settings allocate language to tasks, interactions and locations. One key motivation for this work is to contest the ways in which speakers of non-prestige languages are marginalised by their lack of particular languages. Unfortunately, research and policy-making that seek to redress these inequalities can have unintended side-effects. What should we do about women in labour who cannot always access interpreters when quick decisions about medical intervention are required? Raising the bar of citizenship by imposing language tests on people with limited access to education seems like misdirected effort in solving challenges of this type.

Another approach to the inequities associated with different levels of access to prestigious language varieties is to move from counting the languages spoken by particular people or in particular places to exploring, instead, all the different varieties of language individuals use in the various places in which they seek to establish their identities and communicate with others. As Pennycook and Otsuji put it, 'Metrolingualism is neither monolingualism nor multilingualism; it has to do with the dynamic interrelationship between language practices and urban space.'

Returning to our example of labouring women with limited ability to communicate in the dominant national language, a metrolingual approach could support research that explores the role of language in labour and midwifery care. Findings might identify different roles for various languages; perhaps the family language is best for providing comfort and support. With regard to decisions around medical intervention, findings may suggest that women have well thought-out birth plans in relevant languages, as well as support people available to help them communicate how they want their labour to be managed.

There are a number of different ways in which languages in contact are managed by speakers. Ignoring, for the moment, details to do with the degree of bilingualism, the nature of the contact, the size of the linguistic groups and the power relationships between them, we can predict three overriding kinds of contact scenarios in the longer term:

- **language maintenance –** where speakers of each language continue to use their native language from generation to generation

- **language shift** – where speakers abandon one language in favour of another
- **language creation** – where long-term language contact between the two languages gives rise to a new contact language (it might be a pidgin, a creole or a mixed language).

By imagining these three scenarios, we can broadly explore the different patterns of influence that potentially take place when language groups encounter one another. There are also many examples of situations combining aspects of these different outcomes, as we describe below.

12.2 Language maintenance

> Ko taku reo taku ohooho, ko taku reo taku mapihi mauria. [My language is my awakening, my language is the window to my soul.]
>
> Māori proverb

In a language-maintenance situation, especially where languages are maintained over a long period of time, some features will inevitably cross between languages in contact; typically, this will involve aspects of vocabulary first and then possibly structural features. As we described in Chapter 4, **borrowing** is the usual (admittedly rather odd) term given to this process whereby speakers of a language adopt linguistic features from another language – so they maintain their language, although it changes from the incorporation of these borrowed elements. (Of course, when things are borrowed, the understanding is that they will be returned at some later stage – clearly, this isn't the case when the 'loans' are linguistic!)

Diglossia

When we look at languages living in almost symbiotic relationships, a useful term to know is **diglossia**. The creator of the term, Charles Ferguson, originally applied it to the patterned use of formal and colloquial forms of the same language (for example, by speakers of High German and Swiss German, or different varieties of Arabic), but in a later work he extended the term to describe a similar distribution between two different languages.

The key thing identified by Ferguson was the use of different languages and varieties for different functions. He used the labels

H(igh) and L(ow) to distinguish these (the labels are the fallout of their respective formal and informal domains of use):

- H = the 'high' variety, the dominant language, the standardised language, the literary language
- L = the 'low' variety, a minority language, a non-standard dialect, a non-literary language.

In a diglossic situation, the H and L languages/varieties each have distinct roles to play in the daily life of speakers, and Table 12.2 is an example of how these divisions can work.

Table 12.2 Functions of high and low varieties

Situation	**High variety**	**Low variety**
Sermon in mosque	[illegible]	[illegible]
Instructions to waiters	[illegible]	[illegible]
Public speech	[illegible]	[illegible]
Private conversation	[illegible]	[illegible]
News broadcast	[illegible]	[illegible]
Soap opera	[illegible]	[illegible]
Newspaper editorial	[illegible]	[illegible]
Cartoon caption	[illegible]	[illegible]
Poetry	[illegible]	[illegible]
Folk literature	[illegible]	[illegible]

The Anabaptist horse and buggy communities in Ontario and Pennsylvania are good examples of diglossia at work. Speakers are bilingual in Pennsylvania German and English, and the two languages happily coexist, each with its own separate domains of use. Pennsylvania German (in this schema the L-variety) is the language of home and community, and is not usually used in reading and writing. English (the H-variety) is read and written, and is only spoken when dealing with those outside the community. People also have a knowledge of High German. This is not the same as the Modern High German of today but is (archaic) Luther German, the language of the Luther Bible, with influence from these speakers' own dialect. While it is only ever used for religious purposes, and is therefore functionally restricted, High German is held in high esteem as the word of God. People may not be able to converse in it, but they must be able to read it, and children are taught it by their family and at school.

Where bilingualism is associated with conventional differences in the domains of use for particular languages (in other words, where it has the support of diglossia), the situation is usually stable. However, where either language can be used for any purpose (in other words, where bilingualism is not associated with diglossia), it is much more likely that one language will ultimately be used in all contexts and the other language will be lost. It seems that this linguistic ‘compartmentalisation’ is necessary for the languages to coexist over long periods of time.

The more modern groups of Pennsylvania German speakers show how precarious the survival of the language can be. These people started accommodating English in what seemed to be fairly harmless ways and the effect on the community could not have been anticipated. But, as soon as English was allowed to intrude in what were traditionally Pennsylvania German domains (for example, church services), the shift to English in all other contexts was typically swift and complete. Where this has happened only fragments now remain of originally Pennsylvania German domains (for example, in conversations with more conservative members of the community).

Diglossia with an edge

Although our case-study language Mali is the first language spoken by children in Mali households, it is gradually being replaced in the community by Tok Pisin, the lingua franca of many areas of Papua New Guinea. On the surface, it appears as if there is a diglossic situation in the village of Marunga, with Mali being used within families and Tok Pisin the language of public life; however, there are many pressures to adopt Tok Pisin as the language of everyday communication, and it is gradually extending its reach in the community.

The public use of Tok Pisin is the result of the convention of accommodating to the needs of outsiders when they do not know Mali. Marunga is noteworthy in the Mali community for having a school and a health clinic operated by the Catholic Church. Most of the workers in both the school and the clinic are outsiders who have their own mother tongues. Few of them stay in Marunga long enough to learn Mali – and indeed their opportunities to do so are limited by the Mali preference for addressing them in Tok Pisin. Many topics that are now central to daily life, including schooling, cash cropping and participation in the local political process, also privilege Tok Pisin over Mali. As a result, Mali has not been developed to describe these relatively new domains of activities. Because there are no Mali words for many roles in the local government, there are significant barriers to talking about it in the language. This pattern represents something of a vicious circle in terms of language maintenance since the language is failing to keep pace with changes in the culture of the community, and outsiders have limited opportunities to learn Mali.

In other words, the public/private diglossia that is most readily observable is under pressure, and Tok Pisin is extending into more and more family conversations, displacing the role of Mali.

Translanguaging

As well as devaluing languages other than English, the monolingual mindset also problematises a range of behaviours associated with multilingualism, including **translanguaging** practices. Translanguaging can be defined as a set of multilingual practices in which speakers make use of all their linguistic resources as part of a single integrated system of thinking and communicating. Similarly to the orientation noted above for metrolingualism, researchers on translanguaging argue that these practices are essential to identity in transnational contexts. They emphasise the importance of recognising the ways in which speakers use all their language resources in a moment-by-moment project of communicating identity and meaning in specific social contexts. For the majority of people around the world, and for individuals enacting transnational identities in particular, multilingualism is a fact of life rather than a choice. Away from the monolingual mindset, translanguaging is an unremarkable practice, critical to maintaining relationships across multiple roles and settings.

Translanguaging has become a hot topic in education where language differentiation has traditionally been seen as a critical aspect of developing language ability (see also our discussion of bilingualism in Chapter 17). However, researchers such as Ofelia García now argue that insisting on these boundaries means it is impossible to make a fair assessment of students' learning. García and her colleagues argue that the delineation of languages, particularly by processes such as standardisation (which we discuss in Chapter 15) is a social process and does not map tidily onto the actual language practices of individuals – this is especially the case with bilingual and multilingual people, but their divergence from socially endorsed norms can be understood as evidence of greater, rather than reduced, capacities.

Code-switching

In monolingual speech communities, people choose from a vast repertoire of different linguistic forms and their stylistic choices are tuned to produce just the impression they wish to create. As we saw in Chapter 10, a variety of stylistic choices are possible in a language – not only lexical choices, although these are the most obvious, but also grammar, pronunciation, punctuation and paralinguistic features, such as gesture and facial expression. In multilingual parts of the world, speakers can add to this repertoire and achieve a range of sociolinguistic functions by simply switching between languages. We have already explored these practices of moving between languages in terms of metrolingualism and translanguages; here we consider what happens linguistically during these transitions.

Code-switching (sometimes called ‘language-switching’ or ‘code-shifting’) covers situations where bilingual speakers swap back and forth between languages in a single interaction. The term *translanguaging* directs our attention to the communicative aspects of moving among available language systems whereas the term *code-switching* focuses on the structural properties of these transitions. Attitudes to code-switching vary. In some communities, it is a regularly used, highly valued practice; in others, it is viewed as sloppy or even defective. This reflects a range of beliefs about the status of each language and the legitimacy of mixing languages together. For example, the same person might be happy to mix two low varieties but feel that it is important to keep a high variety pure.

For the villages in the Dhonsiri valley in India, code-switching is a fact of life. Stephen Morey provides an example (shown in example 12.1) of a speaker from Balipather, called Ji Kut, whose parents represent both the Tai Aton and Turung communities. This speaker grew up speaking Aiton but married a Turung speaker and now lives in a Turung village. The example begins in Turung but switches to Tai Aiton (forms marked with two asterisks) with a brief section in Assamese (forms marked with one asterisk).

(12.1)

[illegible]

[illegible]

[illegible]

[illegible]

[illegible] …

[illegible]

‘And we will wipe it, and clean all the dirt and filth away so that it will be gone.’

A number of different factors guide the selection of each language during code-switching. As we have seen, all sorts of different social meanings are attached to languages, and when speakers swap between them, they may be expressing different types of relationships or identities. Certain topics may also be more appropriate in one language rather than another, and the switch itself can be a means of highlighting something in the conversation. In bilingual classrooms, teachers can exploit this by switching languages for important concepts or to gain the attention of their students. Code-switching is a natural speech practice that demonstrates high levels of social and linguistic skill. It is a huge topic that has attracted a lot of attention from linguists, neurolinguists, psychologists, educators and speech scientists. These researchers are interested in all sorts of aspects of the interaction – when and why speakers choose one language over another, how the switching works linguistically (for example, the grammatical patterns involved), and what psycholinguistic processes take place in the code-switching brain.

Some linguists also use ‘code-switching’ to embrace bidialectal speakers. It can therefore refer to switching between language varieties (even closely related ones) – for example, between the standard language and a regional dialect. To complicate matters, ‘code-switching’ is occasionally used even more broadly to accommodate changing between registers (e.g. between formal and more informal versions of the same language). True, the meanings and functions of these practices are similar, but we would rather retain the term ‘style-shifting’ to cover situations where speakers alternate between the styles or registers of one language.

Structural borrowing

As we noted in Chapter 4, all parts of the language can end up being borrowed, even grammatical patterns – although usually this requires long-term and widespread bilingualism between donors and borrowers. In fact, we can imagine a kind of hierarchy of ‘borrowability’. Lexical features are the most easily transferred, and these will be content morphemes (especially nouns and verbs with real-world meanings). Contact doesn’t even have to be that extensive for lexical borrowing to occur. Structural borrowing (e.g. grammatical morphemes, features of syntax and phonological features) is much rarer, though, and requires more intense contact and favourable social factors. Basically, the more grammatical the element, the more contact is required to support any borrowing. Free-standing grammatical forms (such as conjunctions and prepositions) are borrowed more easily than bound forms (such as a past-tense suffix or plural suffix), but if one language borrows heavily enough from the lexicon of another, it will also acquire some derivational and even inflectional affixes along the way – and perhaps phonological features as well (though this is unusual, especially beyond their use in the loan words themselves). Generally, the extent and nature of this kind of borrowing will depend a lot on how intense the interaction is between the languages, and also on the compatibility of the two systems – it is easier to borrow constructions that are similar to existing ones (or at least don't conflict with the existing grammar). Crucial social and psychological factors (beyond simple questions of prestige) also play a role in determining what kinds of linguistic features are involved and the extent of the borrowing.

Borrowing usually involves some sort of gain: the borrowers profit in some way from the transaction. The gain might be social – speakers might take on borrowed words, and even pronunciations, for reasons of fashion or social status. Cross-linguistic influences (or **contact-induced changes** or transfers, as they are sometimes called) might also provide useful additional resources for the language. In the case of vocabulary, for example, the transferred material might replace an expression that has become obsolete or has lost its expressive force. It may simply be that a borrowed word fills a lexical gap.

Pennsylvania German has existed in close contact with English for nearly 400 years, and influence from English has been considerable over this time, not just in the matter of lexical borrowings, but also in the structural aspects of the language. Example 12.2 illustrates the passive construction with a number of English influences.

(12.2)

[illegible]
[illegible]
'The stocking is/was darned by mother.'

These influences are as follows.

- The *bei* 'by' phrase is after the verb *geschtoppt* and not before (the expected German word order).
- The preposition *bei* has now replaced the original *vun* 'of'.
- The auxiliary *sei* 'to be' has taken over from the auxiliary

warre/waerre ‘to become’.

The language has also developed a ‘get’ passive equivalent to the English ‘the stocking got darned’ (using the auxiliary verb *kriegen* ‘get’).

There are many other examples where Pennsylvania German constructions appear to be converging with those of English, and the closeness of the linguistic systems and structural compatibility would be assisting here. In truth, though, most of the changes probably have multiple origins, with different internal and external factors playing a role. Here we have two genetically related languages that, because of their common origin, are travelling in similar directions – in other words, the seeds for many of these changes would have been sown long ago in the parent language. The Pennsylvania German speakers themselves may also have initiated some changes independently, but since most of the developments illustrate well-trodden paths of grammatical change (for example, the ‘get’ passive), they are also likely products of universal cognitive processes driving language use. Finally, there is the issue of English contact. Contact might be the trigger or simply an accelerating influence – it is impossible to know for sure.

Areal linguistics

Where languages occur in close proximity, and there is extensive and stable bilingualism, you can have a situation where structural features spread throughout the area in an even more spectacular fashion than in the English–Pennsylvania German contact situation. In its most extreme form (and this is widely reported around the world), this kind of linguistic **convergence** can occur between languages that are unrelated. Even languages with very different syntactic structures can gradually become structurally similar. In other words, regions containing languages that started off being genetically diverse end up being typologically harmonious. In these cases, the borrowing is mutual – features are shared by the converging languages and so they become very similar in structure. One of the most famous examples is the Balkan Convergence area, or Balkan Sprachbund, involving Romanian, Bulgarian, Serbian, Croatian, Macedonian, Albanian and Greek. As is typical of convergence, there is some transfer of lexical items, but also the spread of grammatical patterns that cannot come from a single Balkan source. These include:

- definite articles that follow rather than precede nouns
- a complex future construction from verbs of volition (such as

'want')

- analytic comparatives where the grammatical meaning is captured by a free morpheme (compare something like *more lovely* in English)
- loss of the infinitive.

The issue of the relationship between contact and genetic affiliation remains a controversial area in linguistics, but we hope that this chapter has been able to convey the extent to which all languages are intertwined with other languages to some degree.

12.3 Language shift

> When a language dies, so much is lost … [L]anguage is the repository of the history of a people. It is their identity.
>
> David Crystal (2003), *English as a Global Language*

Language shift involves changing patterns of language use within a multilingual society. It can be partial (the roles of languages change and shrink or grow), or it can be complete (one language ends up no longer being used at all). More specifically, we can identify three general categories of language shift.

1. *Complete language shift*. The entire community switches to another (usually the dominant) language. In migrant communities, the original language is generally still spoken in the country of origin. In Indigenous communities, this may lead to the loss of the original language.
2. *Partial language shift: domain restriction*. The dominant language of the wider community encroaches on the domains in which the minority language was formerly used. The minority language continues to be spoken in one or two domains.
3. *Partial language shift: population reduction*. The proportion of people who speak the minority language (whether as monolinguals or bilinguals) is markedly reduced in relation to people who only speak the dominant language.

Complete language shift is sometimes called **language death**; however, care needs to be taken when using this term. As commentators repeatedly point out, a language is not a biological creature, and it neither lives nor dies; rather languages are social constructs used by people in communication and people may change the language they use to communicate for various reasons. As we will see below, communities that lose languages are increasingly making strenuous efforts to revive them. These initiatives raise all sorts of interesting questions for linguists about language contact, and challenge our assumptions based on the family tree and the wave models of language change, as discussed in Chapter 11. More generally, the role of continuity in the history of languages is questioned.

Contact-induced language change: Long-term patterns of shift

To illustrate the kinds of changes that can take place in language-contact situations, consider the early contact situation between English and Celtic. English was introduced into Britain in the fifth century ce by various Germanic tribes (the Saxons, the Jutes and the infamous Angles, who gave their names to the place, England, and its language, English). These invaders pushed the native Celts northwards and southwards, and the Celtic languages have been receding ever since (Cornish and Manx disappeared relatively early, though revival movements have been attempting to resuscitate them). The Celts seem to have learnt the language of their Germanic conquerors thoroughly, and the influence from Celtic on English is very slight – confined to place names (for example, Devon and London) and some cultural borrowings like the *bannock* (a kind of scone). Despite the widespread destruction of the Celtic languages, there remained a kind of Celtic underlay that gave rise to the very distinctive pronunciation and intonation you still hear in the Celtic Englishes of various parts of Britain (we describe why you get these phonological effects in a moment). This underlay is called a **sub-stratum**.

Contact-induced changes in such situations of shift are not usually referred to as borrowing, but rather **sub-stratum influences**; elements from the mother tongue of a minority group are transmitted to later generations of speakers of the dominant language. So, the first-language habits (in this case Celtic) are transferred into the second language (in this case English). This cross-linguistic transfer is sometimes given the rather judgemental label of ‘improper learning’ (the idea being that the population imperfectly learns the second language that then comes to dominate).

As in the English–Celtic contact, sub-stratum influences typically involve the transfer of phonological features, and might also involve grammatical patterns. They are different from changes that occur when the population continues to speak its native language (that is, language-maintenance scenarios, where there is no intrusive language pushing out another); as described earlier, in this situation, lexical aspects are incorporated first and foremost, and (in cases of close contact) possibly also structural features. Consider the impact of Aboriginal and Torres Strait Islander languages on Australian English; as in Celtic–English contact, there has been little influence except for the transfer of proper names (around one third of Australia's place names) and terms for cultural items (e.g. *boomerang, corroboree* from the Dharug language).

The patterns of influence we describe here are predictable when we take into account a child's language experience and the order in which different linguistic levels are acquired. As you will learn in Chapter 17, intonation and stress features are the first to be acquired and are therefore the most deeply secured features. Together with phonemes, these features are acquired during a child's first three years. The acquisition of grammatical features continues until much later, during the early school years, and vocabulary, of course, continues to be acquired throughout one's lifetime. It is not surprising, therefore, that the most deeply anchored speech habits are the ones we find most readily transferred from a first to second language in a language-shift scenario. Sometimes, a distinctive intonation is all that remains of a previous bilingual speech community. And, to consider the other side of the coin, it is not surprising that these deeply embedded phonological features are also the ones that show the greatest resistance to influence in a maintained language.

Pennsylvania German → English

As described earlier, the Pennsylvania German–English contact situation is currently a stable bilingual situation, yet there are also signs of potential or actual shift to English. The transfer of Pennsylvania German (the first language) to the English of Pennsylvania German speakers (the second language) involves sub-stratum interference – in other words, the transferral of language habits into a second language. The least amount of influence occurs at the lexical level; the few lexical transfers are confined to untranslatable concepts, such as *Freindschaft* (meaning something like 'family, relations'). The most obvious influence occurs at the prosodic level, in this case a distinctive rise-rise-fall pattern for a range of sentence types, including questions and statements. There are also transfers of consonant and vowel features, such as [dʒ] becoming [tʃ] and [ʌ] becoming [ɔ] (for example, *just* is pronounced as [tʃɔst]).

There is little in the way of grammatical transfers. Since both types of contact – borrowing and sub-stratum-induced influences – can result in grammatical changes, why do we find grammatical transfers from English to Pennsylvania German (as earlier described) but not the other way around? The explanation for this lies in aspects of the childhood-acquisition experience. During the early school years, children would not yet have stabilised their Pennsylvania German grammatical structures, and they would be vulnerable to English influences at this time. The way English is acquired in the formal school setting is significant here. There is considerable prescriptivism, with a strong emphasis on 'proper English'; moreover, outside of school, the children's experience of English is within formal contexts only; that is, in contact with outsiders and as a written language (recall the diglossia of this speech community). All these factors work against the transference of Pennsylvania German grammar into the English of these children.

Language maintenance and reclamation

> Ut ma Malichēna da vandi amēr gēlēm ut nēvēt autha methomon da sok uthi sangar mamēr pētha ngēt ma tamon nanēgt mermer. [We Mali people must be proud of our language and we must really hold on to it and speak it forever.]
>
> Julius Tayul (2002), *Asēk dēchama mambu nēva bang [Stories and Songs from the Village]*

Over the past century, there has been a growing trend for multilingual speakers of small or minority languages to abandon them in favour of the dominant language of the wider community or the nation-state. This has led to a great deal of concern among linguists. Because each language is an embodiment of the world knowledge and cultural practices of its speakers, the loss of languages represents an irreplaceable loss of human intellectual capital. This situation causes a good deal of pain for those who are no longer able to speak their ancestral languages – a fact they strongly associate with their other life experiences such as cultural loss, political and economic oppression, and community breakdown.

In Australia, the effect of English on Aboriginal and Torres Strait Islander languages has been devastating. Before English speakers settled in Australia, the country was remarkably diverse linguistically. Table 12.3 shows the approximate numbers of Aboriginal dialects, languages and language families at the time of early European contact.

Table 12.3 Number of Aboriginal languages spoken in Australia at the time of

early European contact

Language families	26–29
Languages	200–250
Dialects	500–700

For Australia's Aboriginal and Torres Strait Islander communities, language contact and the accompanying imposition of colonisation have brought about a devastating reduction in the use of these languages. Consider the recent figures shown in Table 12.4.

Table 12.4 Findings from the first and second National Indigenous Languages Surveys

Language status	**2005**	**2014**	**Change**
Languages still spoken	145	120	25 languages – drop of 17%
Severely or critically endangered	110	100	10 languages – drop of 9%
Spoken by all age groups	18	13	5 languages – drop of 28%

As the table shows, only around 120 of the original 200–250 languages are actively spoken today, according to the National Indigenous Languages Survey published in 2014. Even the remaining robust languages are under threat, despite vigorous efforts being made to maintain them. Only thirteen have more than 500 speakers, and five languages fell out of this category in the decade between 2004 and 2014. It has been estimated that the number of surviving languages might decline by as much as 50 per cent in the next 20–30 years, as the most critically endangered languages lose their last speakers.

What we lose when languages disappear

In a book entitled *Dying Words*, Australian linguist Nick Evans opens our eyes to other aspects of what we lose when languages disappear. His focus is very much on, as he puts it, the 'looming collapse of human ways of knowing', going beyond the languages themselves to take in evidence from psychology, poetics and even music.

The way we talk sensitises us to certain aspects of the world around us – as Evans puts it:

> Language differences act … as the cookie-cutter shaping the categories we use from the dough of experience … They guide our attention to different facets of reality … They nudge our reasoning in

> different directions.

In an interview with Maria Zijlstra for ABC Radio's *Lingua Franca*, Evans gives an illustration from the Aboriginal language Kayardild and its attention to compass direction:

> If we were conducting this interview in Kayardild and you just told me, 'Move the mic a little closer', or something, a Kayardild instruction would have translated as something like, 'Move the mic slightly to the south', even if it was half an inch. And even though we're sitting in a studio that's dark, we can't see the sun, a Kayardild speaker would know exactly what the layout was because you can't say anything without specifying where people are …
>
> And another time I had to put Pat Gabori, who I begin the book with, into a plane. He was an old man and he was blind. He went blind before he ever saw a car, so he didn't quite understand how you move and sit in seats and so on, and I had to negotiate with him to get him into a small plane. The right thing to say to him was 'Rayinmunkiriliija', which is, 'Enter the plane arse-backwards from the south.' To enter 'arse-backwards from the south', that's a single word in Kayardild. So the language is just shot through with instructions like this, and to speak it you … can't speak Kayardild until you go through the mental rewiring process of deciding to pay attention to space at every point. Speakers, especially if they're boy scouts or orienteers or something, they might know how to do that at particular times, but to live among people where everyone is aware of that at every moment, that's an entirely different experience.

It's not just Australian and Celtic languages that are disappearing at such a distressing speed – Native American, Austronesian and Amazonian languages, to name just a few, are also being lost at a great rate. Language endangerment and language maintenance represent the most urgent challenges confronting the field of linguistics. Many linguists are now involved in some way in fieldwork, language documentation and language revitalisation. Even in those places where local languages have healthy numbers, the high status of languages like English suggests their survival is precarious. According to recent estimates, this century will see more than half of the world's roughly 7000 languages disappear. English is not entirely to blame; the spread of other dominant languages, such as German, French and Spanish, and of non-Indo-European languages such as Mandarin Chinese, Hindi and Indonesian, has also been at the expense of many local languages.

There are currently language communities all over the world crying out for support in helping to maintain their language in the face of probable extinction. Linguists need to know what factors promote language survival and what factors contribute to language loss. How can we best assist communities to document and maintain their traditional languages? Over and over again, we see language spread leading to language shift and then language loss. In Australia and New Zealand, vigorous efforts are now being made to reverse this trend for Indigenous languages. In Chapter 2, we mentioned that New Zealand *Te Kohanga Reo* (language nests for children) and *Kura Kaupapa Māori* (Māori language-medium schools) have had promising results. They are being used as models for revitalising movements in other places in the world. Time will tell just how successful programs like these are in reversing the overall trend towards language loss.

Language loss is not a choice but language revitalisation can be

Language loss is only one of the destructive effects that culture loss has had for Indigenous communities around the world. Early encounters with Europeans typically resulted in devastating waves of illnesses and death, followed by the loss of land and economic freedom. As psychologists Michael Salzman and Michael Halloran describe, these communities experience cultural trauma ‘as the product of a legacy of chronic trauma and unresolved grief across generations originating from stunning loss of lives, lands and vital aspects of Native cultures resulting from contact with Europeans’.

Although linguists and language planners sometimes frame language loss as being a matter of ‘choice’, it is clear that people in these communities are not making free choices but rather responding to the forces of colonisation simply with a view to survival. There are communities all over the Americas, the Pacific and Australia that bear the results of these changes, experiencing high rates of a whole range of social and health-related problems.

Time and again, communities have shown that improved wellbeing results when individuals and families within these communities reconnect to culture, and relearn or, at least, learn about their languages. Language revitalisation or revival can take many forms, and numerous strategies exist that people can use to retrieve information about their heritage language and restore it to their community. They range from formal classes incorporated into the program at local schools and visits to special places, to online resources such as dictionaries and games. Sometimes, the linguistic focus on models of language change that emphasise continuity means we can get bogged down in worrying about how the language revived in the present can never be exactly the same as it was in the past. Communities are justifiably incredibly proud of every bit of cultural and linguistic knowledge and practice that they are able to retrieve. While they, too, want to respect the past and carry forward whatever they can, they are also committed to healing in the present and moving into stronger futures.

Factors contributing to language shift

The following interconnecting factors are usually identified as having an impact in situations of language shift.

Migration

People often move to new areas because of economic opportunities or to escape oppression, and there they might encounter a new language. Unless the numbers of migrants with a single language is overwhelming (as it was with the colonisation of Australia and New Zealand), the language(s) of the migrating group(s) are usually abandoned at some point in favour of the dominant language in the new community. The shift to the dominant language generally takes a few generations and may be favoured or resisted by the group concerned. Different generations of speakers often have different attitudes towards the language, depending on where they were born and how strongly they identify with the original culture. Michael Clyne did interesting work on the shift to English among post-World War II migrants to Australia and noted that the Dutch were quick to shift while other groups were much slower.

Economic change

Industrialisation is a common trigger for large-scale migration, although many other types of commercial, industrial or environmental changes have a profound impact on the viability of small languages. One of the effects of colonisation in Australia and New Zealand, for example, was to radically alter the economic independence of Aboriginal and Torres Strait Islander and Māori communities, particularly through the loss of access to land and the alteration of land productivity, resulting from introduced land-use practices such as grazing of hooved animals (cattle and sheep).

Māori English hybrid compounds: Reconciling identities?

When an English-language newspaper such as the *New Zealand Herald* starts incorporating increasing numbers of Māori words – not just as exemplars of local colour but as workaday terms integrated in more complex expressions – it seems to indicate that things are getting better as far as race relations are concerned. Even so, the new compounds developed using Māori terms can have powerful things to say about the state of New Zealand society. Consider the following forms, reported by Degani and Onysko, which serve to highlight differences between groups even as it seems to unite them. (Note that *Pākēha* is the Māori word for New Zealanders of European descent; the diacritics you see here have been lost in the process of borrowing these terms into English).

> Maori/Pakeha culture, Maori/Pakeha world, Maori/Pakeha land, Maori/Pakeha way, Maori/Pakeha settlement, Maori/Pakeha community, Maori/Pakeha perspective, Maori/Pakeha society, Maori/Pakeha story, Maori/Pakeha heritage, Maori/Pakeha tradition, Maori/Pakeha descent, Maori/Pakeha value, Maori/Pakeha identity, Maori/Pakeha owner, Maori/Pakeha support.

Compare this with the *Kiwi*- compounds:

> Kiwi pair, Kiwi coach, Kiwi girl, Kiwi male, Kiwi soldier, Kiwi artist, Kiwi boy, Kiwi crew, Kiwi customer, Kiwi athlete, Kiwi consumer, Kiwi musician, Kiwi trainer, Kiwi worker, Kiwi buyer, Kiwi audience, Kiwi captain, Kiwi character, Kiwi director, Kiwi duo, Kiwi entrepreneur, Kiwi expats, Kiwi hero, Kiwi mum, Kiwi New Zealander, Kiwi parents, Kiwi star, Kiwi talent, Kiwi businessman, Kiwi chick, Kiwi designer, Kiwi favourite, Kiwi friend, Kiwi mate, Kiwi population, Kiwi rider, Kiwi umpire.

Urbanisation

The movement of large sections of rural populations into urban centres in search of work has been made possible by improvements in transportation and communication. It has also been made worthwhile by industrialisation. This movement brings together large numbers of people from diverse linguistic backgrounds and separates them from their own language communities. This is a significant reason for the widespread use of Tok Pisin in Papua New Guinea at the expense of local languages.

Education

Schooling and universal literacy are widely viewed as avenues to the advantages associated with industrialisation and urbanisation. For most of the nineteenth and twentieth centuries, formal education was used as an overt tool of language oppression, forcing children from linguistic minorities to shift to the dominant language. Even today, in the few cases where monolingual schooling is provided in the minority language, a shift to the dominant language is nevertheless likely, since people generally need access to the social supports associated with the dominant language (for example, healthcare). The best way to achieve maintenance of both languages is through bilingual education that provides opportunities for children in both languages. However, as the experience of many bilingual programs for Aboriginal languages within Australia has shown, it is impossible to overcome the prejudices and disadvantages imposed on a group simply through bilingual education. Unless the other pressures are also addressed, bilingual education can only have a limited impact – and indeed it can be difficult to maintain these programs when the languages they promote are not valued by the wider voting public.

Language loyalties – gains and losses

At the very start of this book, we reflected on the fact that language is a central tool for communicating our humanity to each other. Buckets of ink have rightly been spilled on the great cultural and scientific losses that accrue when a language is no longer spoken, but there are more personal and nuanced costs to these changes of allegiance too. In the poem ‘A Farewell to English, Part Seven’, Michael Hartnett expresses the tensions he felt in an Irish writer having to write in the English language. When it was published in 1975, Hartnett used this volume of poems to announce that he was going to move to writing only in Irish. This decision was controversial, commenting as it did on the role of Irish and English in public life, the arts and, by implication, politics. Ten years later, Hartnett returned to also writing in English. Again, this move was controversial but Hartnett argued that one language was not enough for all that he wanted to express.

How different factors work together

The Pennsylvania German–English contact situation is again instructive here for helping us to identify the complex sociocultural and linguistic conditions under which language maintenance is likely to occur. As described earlier, the norm for a minority language in a prolonged contact situation is eventually language shift – it might be over a few generations or a few hundred years. In this case, however, no shift has taken place in either Canada or the United States. Here, we have an instance of a non-standard, non-written language that has no official recognition (to the extent that it is banned in the school playground in some areas), yet Pennsylvania German has been holding its own well and truly against English for nearly 400 years.

Among the most religiously conservative groups of Anabaptists in North America, the language is the most robust. In this intricate language-culture-ethnicity-religion complex, language proficiency in Pennsylvania German accords with a degree of religious conservatism. Religion and way of life are intimately connected for these groups, and strongly govern their attitude towards their German dialect. Language has a deeply religious significance and this will guarantee its survival. For these people, the question of formal-maintenance efforts never arises. There is no chance of them ever actively propagating the language. For them, it is simple: language and faith are one – one cannot continue without the other. The more modern groups of speakers, however, show the dangers of living next door to a bulldozer language. Even competent speakers are not supported by diglossia and, as discussed earlier, the situation is one that facilitates easy spread and ultimately language shift to English. Among these more progressive groups, the crucial thing is that the children of even the most proficient speakers will not be teaching their own children Pennsylvania German. For these groups of people, loss of Pennsylvania German is the most probable outcome.

Bringing Māori back

Although examples are not widespread, it is important to note that language shift can work in favour of minority language acceptance just as well as majority language dominance. Having dropped to very low levels of use following decades of oppression through the early and mid-twentieth century, Māori is now making a comeback. Language revitalisation efforts relating to Māori got underway in the 1970s and, after decades of community and political activism, language learning and language development, the tide in language attitudes has turned. The music industry is one way of identifying acceptance of Māori: *te reo* (Māori language) albums and songs have gained enormous popularity across New Zealand in recent times. The use of Māori by public figures is also significant. Think of Prime Minister Jacinda Ardern speaking Māori at Buckingham Palace (19 April 2018). And when she announced that her baby daughter, Neve Te Aroha, would be raised learning both English and Māori, she is recorded as saying, ‘We haven’t just made that decision though about how that will happen … It’s an official language. It builds our understanding of Māori culture as well. For me, language is what sits at the heart of that.’

12.4 The creation of new languages

We now move to the third kind of contact scenario, where the outcome is most striking. Under certain conditions, interaction between speakers of different languages sometimes triggers the birth of a brand new language – a pidgin or a creole, or perhaps another type of mixed language.

Pidgin languages

Generally speaking, a pidgin is a type of makeshift language that springs up when speakers of different linguistic backgrounds come into contact and there is an urgent need to talk. In the formation of a pidgin, there are always two (or more) languages that are involved, although the pidgin takes one language, usually the socially dominant one, as its point of origin for the lexicon. This language contributes most of the vocabulary, though significant features of the grammar are likely to derive from other sources.

Traditionally, people thought of pidgins as simply debased forms of whatever the dominant language was, and they used derogatory labels such as ‘broken English’, ‘impoverished English’ and even ‘lingual bastardisation’, ‘monkey-talk’, and ‘gibberish’. This ‘me Tarzan – you Jane’ image for pidgins is totally inaccurate. Pidgins quite clearly have their own grammars; they are not simply defective varieties of another language; however, they do differ from other types of language in a number of respects.

- Pidgins aren't anybody's first language, but are used when speakers venture outside their usual speech community.
- Because they are used in limited contexts, pidgins lack the range of stylistic variation characteristic of a normal language.
- Pidgins make do with reduced vocabularies and require less complex and flexible structures. English-based pidgins, for example, are syntactically and morphologically simpler than the variety of English on which they are based.

Clearly, both Australia and New Zealand offer situations where English has come into close contact with other languages. Since European contact, the Aboriginal and Torres Strait Islander peoples of Australia have seen members of several language groups living in the same community and engaging in daily interaction. Not surprisingly, pidgins based on English appeared not long after the arrival of the Europeans.

In the early 1880s, a 'foreigner-talk' system known as South Seas Jargon was used in various parts of the Pacific, primarily between European whalers and Indigenous crew members, some of whom were Māori (*jargon* here refers to a kind of pre-pidgin hybrid language and is different from our use of jargon in Chapter 10). In New Zealand, this jargon developed into Māori Pidgin English, which was used for early contact between Māori and Pākehā (or settlers).

Creoles

Typically, a pidgin is short-lived. Speakers often abandon it in favour of the source language or, if the contact ceases, the pidgin will simply die. However, if the situation stabilises and the contact continues, the makeshift pidgin metamorphoses into a fully-fledged and dynamic language. Often this occurs when speakers of a number of different, but related, low-prestige languages come together in the context of an unrelated high-prestige language. English-based creole varieties cluster in three main areas: the Caribbean, West Africa and the West Pacific. They include Costa Rican Creole, Gullah (spoken by African Americans in the far south-east of the United States), Jamaican Creole, Krio (Sierra Leone), Bislama (Vanuatu), Tok Pisin (Papua New Guinea) and Kriol (northern parts of Australia). Interestingly, South Seas Jargon never stabilised enough to become a creole. This is mainly because, in New Zealand, Māori was the only alternate language to English, so there was no need for a lingua franca between Indigenous groups. The historical records also suggest that the most common pattern was for English speakers to learn enough Māori to communicate. As a result, the New Zealand pidgin was short-lived.

Because the developing creole must now serve its speakers in all kinds of settings and circumstances, the changes are typically rapid, especially in vocabulary and grammar. In theory, it is straightforward to say when a pidgin ends and a creole begins. As soon as children in a community are brought up speaking the pidgin as their first language, it becomes a creole. Accordingly, a creole is simply a nativised pidgin. The linguistic reality, however, is another matter – linguistically, it is impossible to say where the boundary lies. Even before a pidgin becomes somebody's first language, it can develop a highly elaborate structure (close to that of a creole), if it is used for a number of different purposes. For this reason, some linguists avoid the labels 'pidgin' and 'creole' and refer to these varieties straightforwardly as '**contact languages**'.

Was medieval English a creole?

Various stages are identified in the development of English: Old English (450–1100), Middle English (1100–1500) and Early New English (1500–1800). Of course, stages like these are arbitrary, and the dates and labels are just convenient ways to describe the evolution of the language. If we had the benefit of some sort of time machine, and could arrange a get-together of speakers of English from, say, the fifteenth and sixteenth centuries (so-called Middle English and Early New English speakers), the differences in their speech wouldn't be any more remarkable than what we witness today between older speakers and teenagers. What we have is a continuum – at no point does Old English stop being Old English and start being Middle English! However, there are linguists who have challenged this idea of 'unbroken continuity' and go so far as to suggest that medieval English was not a modernisation of Old English but rather a creole. They support this idea with evidence such as:

- the content of the lexicon – 85 per cent of the Middle English lexicon is French
- the structure – we see a shift from synthetic structure (grammatical meaning conveyed by inflections) to analytic

structure (grammatical meaning conveyed by separate words and word order)

- the socio-historical context – the small English aristocracy maintained its French ways (for example, in speaking French), with the rest of society continuing to speak English.

Evidence like this is suggestive, but while the idea of a medieval creole is interesting, it is far-fetched. The types of changes that took place between the Old and Middle English periods are well documented for contact situations generally. In short, the language of the medieval poet Geoffrey Chaucer is not a linguistic innovation but simply reflects the contact situation (which also involved Scandinavian). Fallout from language contact is all we need to explain the drastic changes that took place in English at this time.

The negative attitudes associated with pidgins that we noted above also get applied to creoles. For example, many expatriates in Papua New Guinea view Tok Pisin as baby talk and feel that locals are not making an effort when they use Tok Pisin, even among themselves. The significant role played by Tok Pisin in ensuring national harmony (evidenced by accommodation to Tok Pisin in contact situations at the village level as in the example of Mali described earlier) is not apparent and becomes very difficult to see from this perspective. These attitudes make the assertion that English has developed from a creole seem excitingly controversial!

The pidgin varieties spoken in Australia became increasingly important for contact, not only between speakers of Aboriginal languages and English speakers, but also as a lingua franca between speakers of different Aboriginal languages. The disruption that followed the spread of the Europeans across the continent associated with the displacement of people to mission stations, pastoral properties and towns meant Aboriginal people of different linguistic groups were thrown together a great deal. Although there had always been widespread multilingualism among adults, this wasn't adequate to cover communication needs in these new settlements, where children of different linguistic backgrounds were thrown together, and where there was continued uneven interaction between speakers of different Aboriginal languages and English speakers. Pidgins fulfilled a communicative need that quickly became entrenched. Out of these pidgins, creoles evolved in the Kimberley Region, the Roper River area and parts of North Queensland. These various English-based creoles have much in common and are collectively known as Kriol, but they also show some regional differences, which reflect the Aboriginal languages represented in the community where the pidgin originated and also influences from other pidgins and creoles brought into Australia from the outside.

Our brief foray into creole structure will focus on Kriol, an Australian variety of creole spoken by more than 20 000 speakers with regional variants as we noted above. The vocabulary of Kriol is largely English based, although phonologically restructured on account of reduced vowel and consonant inventories. There is also influence from Aboriginal languages; Kriol has lamino-palatal and retroflexed consonants that aren't found in English. Some words are also borrowed directly from Aboriginal languages, and familiar-looking English words can have quite different meanings and functions: for example, *dust* 'to overtake on a road'; *lie* 'pretend' (*We lie–don't look* 'we pretended not to look'). But, it is the grammar of Kriol that differs most from the base language, English.

Contact situations are generally calamitous for inflections (this was the case for English in those early days of contact with Scandinavian and French) so, not surprisingly, contact varieties like Kriol typically display limited morphology for both nouns and verbs. However, as creoles stabilise, they start creating new verbal morphology via the normal processes of language change. For example, Kriol adds *-im* to the end of (most) transitive verbs (verbs that take an object) – *Im bin chakim spia* 'He threw the spear'. (You can probably guess that this suffix comes from the English object pronoun *him*.)

In Chapter 10, we described a kind of cline or continuum of variation in these speech communities, ranging from Aboriginal English at one end to Kriol at the other. Here we list a handful of the most striking grammatical features of Kriol, and the closer a variety is to the creole end of the continuum, the more of these features linguists find. Many of them are shared with other creoles; some are found in other non-standard dialects.

- Unmarked verbs can imply tense, but there are also optional tense markers: *bin* (< *been*) signals past tense (*Im bin gilim me* 'he hit me'), and future tense may be expressed with the pre-verbal marker *gonna/na*, or *gotta* (*All the kid gotta go bogie-bat* 'all the kids will go swimming').
- Aspectual distinctions are variously expressed – for example, the clause final *na* for perfect aspect (expressing completed action) and the suffix *-bad, -bat* or *-obad* (< *'bout, about*) or reduplication for continuous aspect (*Im bin megimbad ginu* 'he

was making a canoe' and *Im bin gray gray* 'He was crying').

- There are many different negative constructions – for example, the auxiliary *gan* 'cannot' or *nomo* (*Yu nomo bin albim me* 'you didn't help me'). Some forms are shared with other non-standard varieties: *'e never even die* (here *never* means 'not' unlike Standard English *never* 'not ever').
- Questions can be formed by rising intonation or by tags such as *ngi* or *intit*; interrogative pronouns include *wanim* 'what', *wijan* 'which' and *waya* 'whereabouts'.
- Nouns in Australian creoles generally do not inflect for plural number or possession; however, plural can be indicated by words such as *mab* or *lad* (*dad lad men* 'those men'), and possession can be marked with *blanga/bla* (< *belong*), as in *Dad san bla mai sista im lib na Sydney* 'my sister's son lives in Sydney'.
- Numerals and adjectives often appear with the suffix *-bala* (a feature of many other creole varieties) – for example, *dubala boi* 'two boys'.
- The most striking feature of Kriol from an English speaker's perspective is the elaborate pronoun system that distinguishes dual and plural, as well as inclusive and exclusive first person (the closest Standard English comes to this distinction is the inclusive pronoun form in *let's* as in *let's go now*). Here are just some examples to illustrate.

yundupala	you two
yupala	you more than two
minyu	me and you
mindupala	the two of us, excluding you
wilat	several of us, including you
mela/mipala	several of us, excluding you

Unserdeutsch – a boarding school creole

There appears to be only one German-based creole in the world, and that is Unserdeutsch (or Rabaul Creole German). As linguist

Péter Maitz describes it, the language had its genesis in German New Guinea in an orphanage for mixed-race children, where Standard German was taught (at that time, the whole Bismarck Archipelago was under German colonial rule). In the dorms, the children developed their own group language – a kind of secret language. So, the creole is highly unusual in that it was a tool of in-group communication (in the school context) instead of out-group communication (say, among slaves on plantations), which is the norm for these contact languages.

Unserdeutsch has some typical creole features such poverty of grammatical markers (inflections) and fixed SV(O) word order. But, like Kriol, it also has its own complexity – for example, in the pronoun system, there is a first-person inclusive (incl)–exclusive (excl) distinction *uns-wi* ('we' pronouns that distinguish whether or not the hearer is included):

[illegible]-en
[illegible]inclusive
'The two of us are talking so lovely'
[illegible]en
[illegible]-exclusive
'We dance when we have music'

Alas, Unserdeutsch will soon be extinct. There are only about 100 (elderly) speakers left, nearly all immigrants to Australia, and living in and around Brisbane and Sydney. If you want more of a taste of this fascinating language, we suggest some places to go in the Further reading section at the end of the chapter.

New mixed languages

The linguistic situation in Kriol-speaking communities in Australia is even more complicated than we have been describing here. In an article called 'Language Landscapes of Children in Remote Australia', linguist Jane Simpson focuses on four Indigenous communities in the north of Australia – although, as she emphasises, their linguistic situation is replicated all around the country. Here is how she describes the complex multilingualism of these communities:

> Many Indigenous communities in remote Australia are multilingual, and often the languages being spoken in the community are rapidly changing. Traditional languages are spoken by some people, but at the same time new languages are being developed based on the interaction of traditional

> languages, English and an English-based creole. These new languages vary along a continuum. At one end, the way of talking is close to the way many people in rural Australia talk. At the other end are mixed languages, in which the structure of the new language contains words and features of several languages. In the middle of the range are varieties of an English-based creole. Children in these multilingual communities grow up in language landscapes that are often undergoing rapid change.

The mixed languages referred to here were created via bilingual speakers' integrated use of Kriol and a traditional language. One of our case-study languages, Gurindji Kriol, is an example of a new mixed language of this kind. Spoken by the Gurindji people of the Victoria River District in the Northern Territory, it emerged during the 1970s and combines the lexicon and structure of Gurindji (the traditional language of the Gurindji people) with Kriol. Felicity Meakins, who has worked extensively on the language, offers the following example to show the linguistic mix (the Gurindji-derived forms are italicised and Kriol forms are in plain font).

(12.3)

[illegible]*-ta-ma*
[illegible]-locative-topic
karu-walija-ngku
child-paucal- ergative
'And the kids had sport at school.'

In this example, the core verb phrase structure, including the past tense marker *bin* and transitive marking *-im*, is derived from Kriol, while the noun phrase structure, including the ergative-case marking -*ngku* (showing the agent) and locative-case marking *-ta* (showing place), is from Gurindji. You can see other Gurindji nominal morphology here – for example, the paucal marker *-walija* 'few' and the discourse clitic *-ma* signalling the topic. As this example shows, Gurindji contributes the noun phrase structure, but the actual nouns are derived from both Gurindji (for example, *karu* 'child') and Kriol/ English (for example, *skul* 'school' and *sport*). This example also shows that nouns and main verbs can come from either Gurindji or Kriol, but there is a structural split between the noun phrase and verb phrase grammar. We have just noted that Gurindji contributes the noun structure (including the case marking), but the verb structure – including auxiliaries – comes from Kriol. Kriol also dominates the Gurindji Kriol lexicon, contributing something like 64 per cent of verbs and 67 per cent of nouns.

A linguistic split between grammar and lexicon (with one language contributing the grammar and the other the lexicon) is typical of mixed languages. Ma'a (spoken in Tanzania) is historically

Cushitic (of the Afro-Asiatic language family), with 50 per cent Cushitic vocabulary, but a structure that is largely Bantu (of the Niger-Congo language family). Media Lengua is another example, with its lexicon from Spanish and its grammar derived from Quechua (a Native South American language spoken primarily in the Andes).

12.5 Language choice and linguistic accommodation

> When people from different classes meet, if they like each other or want to get on, they 'accommodate'; i.e. move towards each other's speech style.
>
> Michael Argyle (1994), *The Psychology of Interpersonal Behaviour*

Although our description of diglossia paints a picture of orderly language choices, our translanguaging and code-switching examples, and particularly our discussion of language shift, indicate that there are many factors at play in selecting the language to use in a multilingual setting. Accommodation theory seeks to explain the fact that people tend to alter their manner of speech during a conversation to accommodate to others. This can affect many areas of language, including accent, loudness, rate and style of speech, vocabulary and even grammatical aspects, such as the length and complexity of utterances. **Accommodation** also applies to speakers of different languages in terms of language choice; it is useful as a way of organising our thinking in relation to language selection in complex settings.

This behaviour is potentially reciprocal – one person may do all the accommodating or both may adapt. It is also dynamic. In a study of conversations lasting for half an hour, researchers found that interpersonal and face-to-face aspects of the exchange gained in significance over intergroup considerations, and the content of the conversation became the most significant determinant of language choice. Speech accommodation is a very natural thing to do. Even babies alter their pitch when they are 'talking' to adults – they use a lower pitch to their fathers and a higher pitch to their mothers.

The different ways in which we adapt our speech are linked to social psychological processes, such as cultural identity, attitudes, perceptions and intercultural relations. The following principles serve as the foundation for the theory:

- *similarity attraction* (speakers makes themselves more attractive to others by minimising differences between them)

- *social exchange* (a social psychological mechanism used by speakers in order to determine the rewards and costs of various actions)
- *causal attribution* (the listener interprets the speaker's choices as indicative of particular motives and intentions)
- *intergroup distinctiveness* (participants consider each other and identify the ways in which they are the same or different).

Speakers might **converge** in order to reduce the distance between them and those with whom they are communicating – since people generally react more favourably to those who are linguistically close, this seems like a pretty sensible thing to do. A now-famous study of accommodation by Nikolas Coupland investigated service exchanges in a Cardiff travel agency; he described how the shop assistant altered her pronunciation to match the pronunciation features of her clients (here the variation ranged between the local Cardiffian accent and Received Pronunciation, and hinged on the education and socioeconomic background of the clients). You can see the benefits of accommodation here, but a lot will depend on the situation. Convergence does bring with it some loss of personal identity and, of course, there is always the risk that the interlocutor doesn't appreciate this new style (for instance, imagine how you might feel if your interlocutor picked up features of your own accent).

Moreover, if people converge too closely to the speech of their interlocutors, it can backfire badly. Over-convergence is conceptually possible because there are norms about how much convergence is appropriate in a given setting. There are social scripts associated with all kinds of settings that may involve certain levels of linguistic accommodation and it is possible to go too far. Another famous study, carried out by Howard Giles and Philip Smith in 1979, found that a Canadian English speaker who converged towards a British English speaker on any one dimension (they tested for rate of speech, pronunciation and message content) was perceived positively, but when the same speaker converged on all three at once they were viewed as being condescending. Over-accommodation (sometimes the convergence is to a stereotype rather than the behaviour of the other person) has been identified in caregiver talk when medical personnel talk down to clients or patients (disabled people, the elderly and the ill are particular targets). You can also see it at work when people talk in non-standard, over-simplified ways to foreigners (known as 'foreigner talk'), or to young children.

There are many different ways of responding to a situation through language choice, and speakers can even choose the opposite strategy to accommodation – in other words, they can accentuate their linguistic distinctiveness and make their language less like that of their audience. **Divergence** tends to arise in settings in which speakers want to protect or confirm their (personal or group) identity. Speakers tend to do this more in stressful situations. And although it can be a way of positively strengthening one's social identity, it does run the risk that listeners will perceive linguistic divergence in terms of impoliteness, as an expression of insult or hostility.

The labels we have used here for these different ways of responding to situations through language choice are the same labels that are applied to the structural changes that can take place in language contact. Recall that we talked earlier about convergence with reference to the levelling out of differences that can occur when languages are in close contact and influence each other over an extended period of time. Alternatively, divergence is when speakers of one language move apart, physically or socially, and their language develops along different pathways. So, we have convergence/divergence that involves accommodation in relation to the behaviour of speakers, or convergence/divergence that involves language contact in relation to the structure of languages. Both situations reflect complex processes involving the interaction of linguistic, sociocultural and also psychological factors.

12.6 Translation and interpreting

> Say what one will of the inadequacy of translation, it remains one of the most important and valuable concerns in the whole of world affairs.
>
> Goethe, letter to Thomas Carlyle, July 1827

Translation involves the conversion of a written (or recorded) text from one language to another. **Interpreting**, on the other hand, is the facilitation of oral or sign-language communication between users of different languages. This can be simultaneous (that is, happening in real time) or it can be carried out consecutively (that is, one party finishes and then the interpreting takes place). So, interpreting and translation are not synonymous, even though people do use the words *translate* and *translation* more generally to refer to the conversion of one language to another.

In the task of translation, linguists have tended to situate themselves as those who can support the translation endeavour through their knowledge of language structure. In this section, we want to give you some idea of the types of issues that linguistics addresses in discussions of translation. Leaving aside big issues to do with the more philosophical questions about translatability of texts across cultures, we concentrate on some of the linguistic challenges of the translation endeavour (and these will more or less also hold for interpreting).

We begin by considering the original (source) text and the translated (target) text, and the relationship between them. There are two questions here.

- When a text is translated, what is the overall goal?
- What are some of the complexities involved in meeting this goal?

Of course, an important factor will always be the compatibility of the languages concerned. Things can become quite tricky when two languages are vastly different from each other, not only grammatically and lexically, but also culturally and socially.

First, there are the lexical and semantic challenges. In our discussion of collocation in Chapter 6, we noted that complex relationships hold between form and meaning, especially when different languages are involved. A single form may have more than one meaning and the meanings collectively will not be the same in different languages. Recall that German people *putzen* (that is, ‘clean’) their teeth, not *brush* them as English speakers do. As discussed in chapter 6, the verb *putzen* can also be used in relation to noses and babies’ bottoms (not *wipe*), shoes (not *polish*) and birds preening their feathers.

Another difficulty lies in the incompatible connotations of what would otherwise seem to be equivalent words. An old example – but a good one – is given by Robert Adams (1973):

> At a climactic moment in Stendhal’s *Le rouge et le noir* (Book II, ch. 19), Julien Sorel, after weeks of solitary suffering, has finally climbed back into Mathilde de la Mole’s good graces, and so undertakes once more the perilous ascent, via a ladder, to her midnight bedroom. She receives him with ecstatic, unbounded delight, crying, ‘C’est donc toi!’ And just here C.K. Scott-Moncrieff – for whose extraordinary gifts as a translator I have, as a general rule, only the highest respect – slips on the insidious banana peel, and translates, ‘So it is thou!’ What girl of high social rank and free social manners ever greeted a lover that way?

While *thou* has historical uses akin to the French use of the second person singular *tu/toi* among intimates, it is semantically quite marked in Modern English. As Adams notes, the word *thou* now has archaic and religious overtones and this is his basis for objecting to the use of the pronoun in the above text. A person who knew a bit of French might be able to make the link the translator intended but, since the intimacy of the situation is more than adequately indicated by the expressions ‘her midnight bedroom’ and ‘She receives him with ecstatic, unbounded delight’, it hardly seems to be necessary.

Another difficult aspect of translation involves the lack of corresponding forms and meanings in the target language – cultural concepts across languages are rarely a perfect match. Over the years, linguists have provided different ways for dealing with words that lack corresponding forms. One strategy might be to use a cultural substitute, as in the following examples: *lamp* = *bamboo torch* (PNG), *corner stone* = *main pole* (of house) (PNG), *lion* = *jaguar* (South America).

Of course, it is not simply vocabulary that can be incompatible across languages. Mismatches between grammars can make life very difficult for translators. Mali makes use of noun class marking as a means of conveying additional information about nouns, creating nuances that are sometimes difficult to render in English. Russian has no definite or indefinite articles (akin to English’s *the, a* and *an*). Japanese has a complex set of sentence-final particles that are impossible to render neatly in English. These are just some of the grammatical challenges, and they can be compounded by differences in pragmatic conventions. As we will see in Chapter 13, discourse markers (such as *well, yeah–no* and *you know*) are highly idiosyncratic features of a language and create a lot of headaches for translators, as do aspects of style.

Since interpreters work with spoken or signed languages, they usually involve a completely different set of professionals (sometimes they are characterised as the extraverts in this field, whereas the translators, who work with written language, are the introverts!). Their role is often described in terms of a channel or bridge (even a kind of human telephone) that enables communication between two people to take place. Interestingly, these metaphors imply interpreters are themselves totally neutral and have no participation in the interaction itself. We could use the rest of this book to debate how realistic, or even desirable, it is to define interpreters in this way. The fact that they communicate across cultures as well as across languages already suggests they have to be more than inactive translation devices.

Interpreting probably has more visibility than translation because of its use in the public arena – most people are familiar with the presence of headphones at major international conferences and high-profile political meetings. But, there is a lot of less visible, but still crucial, interpreting work that goes on in everyday life – in law courts, health clinics, hospitals, tribunals, workshops and seminars, individual or client-group interviews, and so on. Diana Eades has done a considerable amount of work to promote the crucial role of interpreters in the legal system. Courtroom talk has its own challenges, of course – even for people who are fluent in the dominant language – but the difficulties are compounded for second-language speakers. And the nature of the legal process poses additional challenges for interpreters, especially where sign-language users are involved. Some relate to specific linguistic differences between signed languages and spoken languages, but other challenges emerge because sign-language interpreting tends to be simultaneous, rather than consecutive; as Eades describes, this has serious drawbacks in the legal setting (beyond the exhaustion levels of the legal interpreters who have no time to rest or pose questions).

The changing role of Auslan interpreters

Recently, there has been a lot of interesting discussion around the role of interpreters and their work with the Deaf community. There have also been a lot of changes in their role in this community. Here are the ways interpreters for the Deaf have been characterised over the past century:

- *Interpreters as helpers.* Before the 1960s, interpreting was typically done by family and friends in an unofficial capacity. These people viewed themselves as helpers and generally felt free to add in their own contribution.
- *Interpreters as conduits.* In an effort to get away from seeing the Deaf person as helpless in the interaction and to avoid the prospect of the interpreter taking over the decision-making responsibilities of the person communicating ‘through’ them, the interpreter was (and, to an extent, still is) described as a machine that communicates meaning.
- *Interpreters as communication facilitators.* The interpreter is the channel through which the message passes on its way from the sender to the receiver.
- *Interpreters as bilingual and bicultural specialists.* The interpreter is trained to convert utterances between two language systems and also between two cultures.

Something we haven’t considered here (although we will in Chapter 18) is the process and potential of machine and computeraided translation. There are already free online translation sites, one of which is called Babel Fish. It derives its name from the original brainchild of Douglas Adams in his radio series *The Hitchhiker's Guide to the Galaxy* (broadcast on BBC Radio in 1978). Babel Fish was the universal translator that Douglas described as:

> small, yellow, leech-like, and probably the oddest thing in the universe. It feeds on brain wave energy, absorbing all unconscious frequencies and then secreting telepathically a matrix formed from the conscious frequencies and nerve signals picked up from the speech centres of the brain, the practical upshot of which is that if you stick one in your ear, you can instantly understand anything said to you in any form of language: the speech you hear decodes the brain wave matrix.

While a portable language interpreter along the lines of Babel Fish is still a long way off, software technology is advancing rapidly. Will there be a time when machines will be able to translate more accurately and faster that any human second-language speaker?

Further reading

There are some excellent books on contact linguistics, including Clyne (2003), Winford (2003) and Myers-Scotton (2002). On metrolingualism, see Pennycook and Otsuji (2015); our quote is from page 19. On aspects of code-switching, see Myers-Scotton and Jake (2009), and on contact languages see Siegel (2008). For more on the 'monolingual mindset' see Clyne (2008). On language maintenance and language shift in Pennsylvania German-speaking communities, see Burridge (2002, 2007). For diglossia, see Ferguson (1959). On translanguaging, see Otheguy, García and Reid (2015). More information on the languages of Aboriginal and Torres Strait Islander communities is available in AIATSIS and FATSIL (2005), and Marmion, Obata and Troy (2014). Our discussion of language choice and speech accommodation was adapted from Hamers and Blanc (2000), and relied particularly on Coupland et al. (1988). See also Giles and Smith (1979). Larson (1998) provides further discussion on intercultural challenges in translation; see Eades (2010) on the role of the interpreter. The full *Lingua Franca* interview with Nick Evans can be found at the ABC's Lingua Franca web page (http://www.abc.net.au/radionational/programs/archived/linguafranca/dying-words/3133192). On language endangerment, see Evans (2010) and Zuckermann (2013); Salzman and Halloran (2004) describe the relationships between cultural trauma and language loss. See Hartnett (1975) for a collection of his poetry. Discussion on the resurgence of

Māori includes journalism by Ainge Roy (2018), as well as academic work such as that by Degani and Onysko (2010) on the hybrid compounds in New Zealand English. For a complete timeline of the history of Māori, see the Māori Language Commission (Te Taura Whiri i te Reo Māori) website (http://www.tetaurawhiri.govt.nz). Domingue (1977) puts the case for Middle English as a creole. Recent work on Kriol includes a series of papers by Mushin and others (see, for example, Mushin and Watts 2016) and by Dickson (2016). For information on Unserdutsch, see Maitz and Volker (2017); there is a very comprehensive website on the creole hosted by the University of Augsburg, which has links to an Unserdeutsch Facebook page, and a short SBS World News broadcast featuring Péter Maitz (https://www.philhist.uni-augsburg.de/en/lehrstuehle/germanistik/sprachwissenschaft/rabaul_creole_german). Our quotation on mixed languages is taken from the abstract of Simpson's (2008) article (page 51). The discussion on Gurindji Kriol follows Meakins (2013a, 2013b); our quote is from Meakins 2008: 0247.

Exercises and discussion points

1 A closer look at Gurindji Kriol

Gurindji Kriol belongs to a set of contact languages that draw their lexicon and grammar from two languages. Remembering that Kriol is an English-based contributor to Gurindji Kriol, ascertain the extent to which the following examples bear out this characterisation of the language. Nouns and their associated morphology have been marked in bold.

[illegible]*-ngka*
[illegible]-at
'The child is sitting on top of the deer.'
[illegible]
[illegible]-TRANSITIVE
[illegible] ***-ngka***
[illegible]-at
'It (the deer) took the child running on its head, with the dog below.'
[illegible] ***diya****-ngku*
[illegible]-agent
[illegible]*-ma*
[illegible]-towards
'The deer threw the child and the dog downwards off the cliff.'
[illegible]***- parni***
[illegible]

jirrpu	*ngawa-ngkirri*
dive	water-to

‘The pair of them, the child and the dog fell down, diving into the water.’

2 Back-translation – for those who know more than one language

In back-translation, one person translates a text from Language X into Language Y, then a different person translates the Language Y text back into Language X. It is a way to check the quality of the translation (the closer the versions, the better the translation). Something similar can be done to check the quality of a machine translation. Take English and another language you know. Choose a simple text in English (for example, a verse of a nursery rhyme or an advertisement) and translate it into this language using a popular translator site (such as Google Translate or Babel Fish). Then take the new translation and, using the same means, translate it back to English. Identify the errors and try to explain what has gone wrong.

3 Language maintenance

Various community languages other than English (LOTEs) are spoken in Australia. Select one LOTE and investigate its current linguistic situation. Provide a brief history of the LOTE, and discuss the domains of use, language attitudes and any efforts by the community towards maintenance of the language.

4 Life stories and language

Who do you know from among your family and friends who has shifted in the languages they use over their lifetime? Ask a couple of people about which languages they spoke and how, when and why they made changes as their life progressed. If they are comfortable talking about this topic (be aware that there may be some difficult memories associated with some of these changes), you may also want to ask them about how they feel about these changes. Are there languages they particularly miss (if they rarely use them now)? Do they feel that their identity shifts as their language choices change? If you have also moved between languages (or dialects) over your lifetime, what are your experiences? Do you recognise elements of what you are hearing in others’ stories?

5 Speaker numbers

The following list contains the ten most widely spoken languages in the world. For each language, say what factors have supported the widespread use of the language.

Language	**Location**
Mandarin	China, Taiwan and Singapore
Spanish	The Americas, Spain, a couple of parts of Northern Africa and areas of the Pacific
English	The United Kingdom and former British colonies, including Australia, Canada, India, Ireland, New Zealand, United States, South Africa, Singapore, and areas of the Pacific
Hindi	India, Fiji and Nepal
Arabic	North Africa, the Middle East, East Africa
Portuguese	Portugal, Brazil, and former colonies in Africa, the Atlantic and Southeast Asia
Bengali	Bangladesh and India
Russian	Russia and former republics of the Soviet Union as well as Israel
Japanese	Japan
Punjabi	India, Pakistan

(a) Are there any languages on this list that surprise you?

(b) Are there other languages you may have expected to find in this list? Which ones? Why do you think they are not here?

Research project

A comparative account of one of our case-study languages

Select and research any one of our case-study languages, then write an account of the various ways in which this language differs from your own language. Identify any mismatches that would cause difficulties for translation or interpreting. You may include social and cultural aspects here, as well as linguistic features relating to structure (sounds and grammar), or you may decide to focus your description on only one particularly distinctive feature. It might be that your chosen language has, say, very different politeness features, as in the case of Japanese compared with English.

For more activities, please go to www.cambridge.edu.au/academic/linguistics.

Part 5

◆

‘Language that rolls up its sleeves’

Language at work

13

Language, Context and interaction

◆

Man does not live by bread alone: his other necessity is communication.

Charles F. Hockett (1958), *A Course in Modern Linguistics*

13.0 Introduction

In this chapter, we begin to draw together the strands of language structure and language variation that were explored in Chapters 9 and 10. Since sentences rarely occur in isolation, the only way to get the full picture of how language works is to go beyond the construction of sentences to the construction of larger pieces of discourse, such as conversations and narratives. As you read through this chapter, we will move beyond the level of clauses and sentences to consider how language is organised into larger and larger structures that communicate increasingly complex messages. We are looking at the ways in which speakers and writers go about packaging their messages – in other words, how they distribute information throughout a text and provide the right sorts of cues to help their audiences follow and interpret the communication appropriately. (Note that some linguists make a distinction between 'discourse' and 'text', but here we use both terms to refer to any continuous stretch of language – both spoken and written – that is larger than a sentence.) The construction of even everyday discourse is much like that of a story – beginnings and endings, shifts in focus, the arrival of new players and changes of scenes all need to be signalled in some way.

We begin by exploring different strategies for managing information structure and describe different types of complex sentences that do this work in English. We also consider strategies of cohesion. These strategies make a text hang together; they link sentences to each other and to the context, helping them make sense. We note how the structural patterns found in conversation and narrative can vary significantly across different cultures, and close by introducing you to a set of tools that help researchers to learn about the settings in which communication takes place.

You will find that this chapter meshes closely with Chapter 14, which is also concerned with larger units of language structure. However, the emphasis here is on how larger sequences of language are built up, whereas Chapter 14 focuses on how people go about understanding each other's intended meanings.

13.1 Information packaging

Think about the sort of grammatical overkill you find in languages. English speakers and writers, for example, have four different ways of, basically, saying the same thing.

(13.1) Red Riding Hood gave Grandma some healthful snacks.

(13.2) Red Riding Hood gave some healthful snacks to Grandma.

(13.3) Grandma was given some healthful snacks by Red Riding Hood.

(13.4) Some healthful snacks were given to Grandma by Red Riding Hood.

While these four sentences don't contrast in basic (or propositional) meaning, in practice they would occur in very different contexts of use. The construction we choose depends on what part of our message we want to highlight, what we think is most important, what is new and exciting information and what we assume our audience already knows. You might also feel their stylistic effects are different – the last two are passive constructions that are more formal and less likely to occur in everyday spoken language.

If it were assumed that Red Riding Hood gave something to Grandma and the healthful snacks were new information, this context would favour example 13.1; however, a context in which the disposing of the healthful snacks was already known would favour example 13.2. In this latter case, Grandma would be the point where we would expect people to pay special attention – it is usual for more informative material to appear at the end of sentences and for the less informative material to appear up front. If Red Riding Hood were the new arrival in this story, the versions in examples 13.3 and 13.4 would be appropriate.

So, languages have grammatical overkill for a reason. Compare how you select from different vocabulary items (for example, which word might best suit your needs from *dog, pooch, hound* and *canine*?). It's a matter of choice. In the same way, we select those grammatical constructions best suited for getting our meanings across. It's not unlike marketing: the content doesn't change, but effective packaging can make a world of difference when it comes to success or failure in promoting a product – or message!

The different ways in which we highlight information are collectively known as **discourse strategies**. Our focus here will be on grammar – specifically syntax (word order, special constructions) and morphology (such as specific markers of discourse status). But, never lose sight of the fact that these strategies potentially involve all linguistic levels, including the lexicon (that is, expressions such as *as for* and *well*), phonology (especially prosodic features such as intonation and pausing), and elements of body language (for example, gesture and eye contact) – all these are used to achieve the appropriate packaging for the meanings of our messages.

In the rest of this section, we look at various overriding principles that govern the flow of information in texts. We then move to grammatical structures that follow on from these principles (see section 13.2).

Given information comes before new

In writing, and also in speaking, there are two basic kinds of information: given (or old) and new. **Given** information is already familiar to the hearer. It can refer to something that has appeared earlier in the text, or it can be given in the sense that it is common knowledge – perhaps privately or culturally shared. **New** information is what drives the discourse forward. It is where we expect our audience to pay special attention.

In most, perhaps even all, languages, the unmarked organisational principle is to move from given material to new material; that is, less informative to new content. This flow of information (the 'given before new' order) appears to have a value in language processing; it is easier for production and comprehension to have an early mention of things that are already established (and therefore more accessible and easier to talk about) before new or unpredictable information is introduced. Texts that follow the given–new preference appear more cohesive, because the given information at the beginning links the sentence to the previous discourse, while the new information is usually taken up in the subsequent discourse.

This basic organisation means that adult speakers and writers typically place important information towards the end of the sentence (it’s not so clear-cut for children, who seem to play by different rules). The following set of questions and answers from Russian (where word order is flexible) nicely illustrates this general principle.

(13.5)

[illegible]ends Viktor?’
[illegible] defended by Maksim’
[illegible] information]

(13.6)

[illegible] does Maksim defend?’
[illegible] defends Viktor’
[illegible] information]

The answer to example 13.5 has the word order as object–verb–subject and example 13.6 reverses the major players with a subject–verb–object order. The simple reason for placing the subject Maxim last in example 13.5 is that his involvement is what is new and surprising. This is the aspect to which readers are expected to pay special attention, whereas Viktor, and the fact that he is being defended, is taken for granted and so appears early. In the answer to example 13.6, however, Viktor is the new information and therefore appears last.

Now consider how a language like English, where word order is fairly rigid, ensures that old information precedes the new information. All our examples here come from the European folktale ‘(Little) Red Riding Hood’ – not the classic tale, but a tongue-in-cheek version rewritten for more enlightened times by James Finn Garner. Here is the start of the tale:

(13.7) There once was a young person named Red Riding Hood who lived with her mother on the edge of a large wood.

This is the classic story opening and the reason is simple. The beginning of a sentence is reserved for information that has a lower communicative value. Here we are being introduced to Red Riding Hood for the first time, so it is more effective to delay her mention until later in the sentence. The use of *there* achieves this. But, once Red Riding Hood has been presented, she is no longer the surprising part of the message and she can appear early in the next sentence. Here's how the tale continues.

(13.8) One day her mother asked her to take a basket of fresh fruit and mineral water to her grandmother's house.

Red Riding Hood is now established information and is appropriately replaced by pronoun forms (*her*); to keep repeating 'Red Riding Hood' would be unnecessary – in fact, it would sound odd. The story then unfolds and we learn what is required of her. The end parts of sentences always drive a story forward because they contain what is newsworthy and exciting. This is known as the **principle of end-focus**.

The **principle of end-focus** states that the unmarked word order is for new important information to occur at the end of the clause.

Clearly, the length and complexity of the material are tied in with the informational aspects of sentence meaning. Because it needs so much more detail, new information is typically longer than old information, and constituents are more likely to be complex or ‘heavy’ (for instance, weighed down by modifiers), so this heavy stuff is better placed late in a sentence. This is the **principle of end-weight**.

The **principle of end-weight** states that the unmarked ordering of phrases and clauses by size within a sentence is to move from small to large.

Look at the italicised constituent in example 13.9a:

(13.9a) The forest was *a foreboding and dangerous place that no one ever set foot in.*

Now notice what happens when this constituent appears early in the sentence:

(13.9b) A foreboding and dangerous place that no one ever set foot in *was the forest.*

The altered version in (13.9b) is grammatical, but it becomes quite clumsy when the heavy noun phrase appears at the start and is followed by its short and not terribly newsworthy predicate (here in italics). It is structurally unbalanced (top-heavy), and also harder to process – we have to hold on to a lot of information before we reach the end of the sentence.

The end-focus and the end-weight principles work hand in hand. New information requires more detail than old information; the unmarked word order is for longer, more complicated and ‘heavier’ structures to appear towards the end of the clause.

Topic comes before comment

The way coherent discourse usually works is that the speaker or writer expresses a series of thoughts about an entity – perhaps a person or a thing. A useful distinction to draw here is **topic** versus **comment**. The topic (or **theme**, as it is sometimes called) is the part of the sentence that indicates what is being written or talked about. It is the perspective from which a sentence may be viewed. The topic is basically a discourse-oriented notion. Because it characterises what the sentence is about, it is typically given. The rest of the sentence makes some sort of statement about the topic and this is called the **comment** (or **rheme**). In the natural order of things, the comment (providing the new information) follows on from its topic – if the topic is what we are talking or writing about, the comment is what we are saying or writing about it.

Even though topic-comment interacts very closely with the given-new distinction just discussed, there are occasions when they part company, so we need to keep them distinct. Speakers and writers might switch to new topics and, of course, at the very start of a discourse, topics are introduced for the very first time (as in example 13.7).

Let's take a look at how topics and comments are organised in some of our case-study languages.

Auslan

The ordering of elements is significant in the way that Auslan signers manage the flow of information. Johnston and Schembri provide several possible ways to communicate the proposition 'My friend has bought a house in Adelaide', depending on the information status of the constituents and what is being emphasised in the comment of the utterance. They can roughly be glossed as follows and cover all the basic possibilities.

my friend buy house adelaide
adelaide my friend buy house
house my friend buy adelaide
my friend buy what house adelaide
my friend buy house where adelaide
my friend do-what buy house adelaide
buy house adelaide who my friend

Mali

Although English tends to treat the subject and the topic as the same feature, Mali makes use of a slot outside the clause itself for expressing topics, so subjects that are marked as topical are indicated here – see example 13.10. The word *a-chumes-ka* 'the man' is the topic and the clause boundary is marked by the following word *da*, which translates loosely as ‘and’ along with the adverbial *sai-sa* ‘happening again’. The subject of the sentence is the pronoun *kē* ‘he’, which refers back to the topic, *a-chumes-ka* ‘the man’.

(13.10)

[illegible]mes-ka
[illegible]person-masculine.singular
‘The man, he will go back again.’

When the topic needs to be clearly identified – perhaps in contrast to a previous topic – this position is sometimes filled by a special series of pronouns used to express contrastive focus.

(13.11)

[illegible]ki
[illegible]contrastive
‘but the other one, she would stay down on the ground.’

Japanese

Japanese verbs always go to the end of the clause, with subjects usually before objects (though objects can also appear earlier). The language has been described as *topic-prominent*, which means there is a strong tendency to state up front the topic of the sentence, separately from the subject; an example would be *Zō wa hana ga nagaĭ* ‘As for elephants, (the) noses are long’. The marker *wa* indicates the topic (elephants as a generic category), which is different from the subject (their noses) of the sentence. Here is a more detailed example:

(13.12)

[illegible]-wa

fish topic-non-subject red snapper subject delicious
‘Among fish, red snapper is most delicious.’ or ‘Red snapper is the most delicious fish.’

Principle of front focus

We have just seen that ends of sentences are important for communication because they contain what is of special significance. So, in the unmarked order, new information appears in this end position. However, there are also several ways in which information can be put at the front in order to particularly signal its significance. This phenomenon is known as the principle of **front-focus**.

The principle of **front-focus** states that the unmarked position for new information is at the end of the sentence, but to bring new information into special prominence, place it at the start of the sentence.

Doing the unexpected is a good way of getting noticed. By bringing forward material that would usually be placed later, the effect is to give it much greater prominence. It is an effective way of grabbing the hearer’s (or reader’s) attention. In English, for example, subjects normally appear in initial position (ignoring conjunctions and short adverb phrases that can precede). This is the basic (or thematically neutral) word order for English, and arrangements that place something other than the subject in this position are less usual and are more marked. Consider this extract from ‘Red Riding Hood’:

(13.13) Red Riding Hood walked on along the main path.
But, because his status outside society had freed him from slavish adherence to linear, Western-style thought, the Wolf **knew of a quicker route to Grandma’s house**.

In the second sentence, the new information (in bold) is in its usual end position, but a second focus falls on the interesting background information (here italicised). It appears in front position, setting the scene and providing a cohesive link to what has preceded.

In short, both the ends and beginnings of sentences are discourse spotlights, but different things activate them. This may seem like doublespeak, but hopefully all will become clear after the next section, when you will see some of the grammatical strategies at work.

13.2 Special discourse strategies

In order to see just how information packaging works, let's take a slightly longer extract from the start of 'Red Riding Hood':

> **(13.14)** [1] There once was a young person named Red Riding Hood who lived with her mother on the edge of a large wood. [2] One day her mother asked her to take a basket of fresh fruit and mineral water to her grandmother's house – not because this was womyn's work, mind you, but because the deed was generous and helped engender a feeling of community. [3] Furthermore, her grandmother was not sick, but rather was in full physical and mental health and was fully capable of taking care of herself as a mature adult.

As the following table shows, each of the sentences expresses a comment about some entity, and each sentence falls into two parts. The early part relates the sentence to what's gone on earlier in the discourse. The second part is much longer and more newsworthy – it's here we find new information about already established participants in the discourse (Red Riding Hood and members of her family).

Table 13.1 Information packaging – given versus new material

Sentence	Given material (less informative) Background	New material (more informative) Focus	
1	There once was	a young person named *Red Riding Hood* who lived with her mother …	*Red Riding Hood* is introduced as new topic
2	One day *her* mother	asked *her* to take a basket of fresh fruit and mineral water to her grandmother's house …	Old information (introduced in sentence 1)
3	Furthermore, *her grandmother*	was not sick, but rather was in full physical and mental health and was fully capable …	Old information (introduced in sentence 2)

What is introduced as new information then becomes old and is taken up again at the beginning of the next sentence. The result is a kind of chaining effect of old (or ‘givenness’) followed by new (or ‘aboutness’), and this arrangement is what gives the passage its coherence. You can see that there are also some adverbial elements here (such as *one day* and *furthermore*), which appear initially and also help the information flow, either by setting the scene or by linking the sentences and creating cohesion (more on that later).

Unlike languages such as Russian, English doesn’t allow constituents to easily move around in the sentence to reflect their information status. This applies especially to writers – as you’ll see later, speakers are a bit freer in this regard. Instead, there are specific grammatical resources that permit constituents to be reorganised. We will mention just six here. The first three (passives, existential constructions and extraposition) are subject-forming strategies, because in rejigging the order of participants they either: (1) establish new subjects (from noun phrases originally functioning as objects), or (2) create empty (or ‘dummy’) subjects such as *there* and *it*. Note that many of the world’s languages use similar strategies for expressive or communicational ends, but our focus here will be on English.

Passives: Creating new subjects

Look at the sentence below. It shows the neutral subject-initial pattern, with a subject followed by the verb and then the object. This is the **active** version.

(13.15a)

A passing woodchopper person	*heard*	*[her screams]*
Subject	Verb	Object

The **passive** version of this sentence reverses the order. The original object becomes the grammatical subject while the original subject gets moved into a prepositional phrase *by a passing woodchopper-person* (which functions as an adverbial).

(13.15b)

[Her screams]	*were heard*	*by a passing woodchopper-person.*
Subject	Verb	Adverbial

As described in Chapter 9, the basic meaning remains the same (the semantic roles are identical), but the emphasis has changed. The passive is the original version in the fairy tale used here. We already know that Red Riding Hood screams, so her screams comprise the given information and, therefore, the less informative part of the sentence. What readers should be focusing their attention on is the *passing woodchopper-person* (who is, in fact, the villain in this version of the tale) – hence the passive construction allows this constituent to appear in the end position, appropriate to its new information status. Whether the active or passive is used is not a matter of 'grammar', but rather depends on the discourse context, and considerations such as new versus old information, focus and emphasis.

The passive also allows us to omit the agent altogether. This can be because the agent is irrelevant information, because it is already obvious or even because it is undesirable to mention it. The following examples refer to contexts in which responsibility may be difficult to assign (or assume) – for example, because the agent is unknown (example 13.16), already well established and therefore unexciting (example 13.17), already known and not in focus (example 13.18), or hiding behind a committee (example 13.19). The patient subject in each sentence is in italics.

(13.16) *No actual wolves* were harmed in the writing of the story.

(13.17) *The woodchopper-person's head* was cut off.

(13.18) *Red Riding Hood's space* was wilfully invaded.

(13.19) *A decision* was made.

There are other aspects to agentless passives. Something like *a decision was made* not only sets up the impression of a general consensus; its impersonal style also lends a kind of sham authority to the wording – presumably because it smacks of prestigious jargons from the sciences. Some people are good at using this kind of hidden bias in grammar and it starts very early. For example, the child who says 'The window got broken' has learnt the advantage of the agentless passive construction, especially in combination with a *get* auxiliary verb. The construction nicely conceals the agent and, in combination with *get*, shifts responsibility to something or someone else.

Existentials (or there-constructions)

As we saw earlier, sometimes the topic is completely new information that we want to reserve for the end of the sentence. Existential clauses provide a structure that can be used to get around this problem. They involve a dummy subject using the pronoun *there*, which then appears in the position normally occupied by the subject noun phrase. *There* has no lexical meaning. Its purpose is purely grammatical: to hold the place of the understood subject. The use of this dummy pronoun *there* allows the subject to appear at the end of the clause where it is most appropriate from an information-structuring perspective. In other words, the logical (or understood) subject, with all its new and exciting information, appears later, giving it much greater prominence.

(13.20a)

[basic clause]

Some fat-free, sodium-free snacks were in the basket.

(13.20b)

[existential clause]

There were *some fat-free, sodium-free snacks* in the basket.

The reason we call the resulting sentence an **existential** sentence is because it expresses the existence of something and this is often how new topics are introduced. It's the stock-standard fairy-tale opener: 'There once was a young person named Red Riding Hood.'

Sometimes the existential *there* is compulsory (when the clause contains only the subject and verb). Try to create a basic clause from the following:

(13.21) There was a basket of fresh fruit and mineral water.

Extraposition

Extraposition changes a complex sentence around so that a subject clause can appear in the end position (it's called extraposition because the subject clause appears outside its normal position). In place of the original subject, we put the pronoun *it*. This is another dummy pronoun subject – a kind of placeholder that anticipates the information coming up in the following clause.

(13.22a)

[basic sentence] *For a little girl to walk through these woods alone* isn't safe.

(13.22b)

[extraposition] It isn't safe *for a little girl to walk through these woods alone*.

(13.23a)

[basic sentence] *To set up an alternative household based on mutual respect and cooperation* was decided.

(13.23b)

[extraposition] It was decided *to set up an alternative household based on mutual respect and cooperation.*

In the above examples, both the extraposed and non-extraposed versions are possible – although, as we said, the former are preferred. The motivation here is the principle of end-weight. The heavier (and more newsworthy) the clause, the more likely it is that extraposition will occur. There are, however, several verbs for which extraposition is mandatory – for example, *seem, appear* and *happen.*

(13.24a) It just so happened *that Red Riding Hood, Grandma, and the Wolf now felt a certain commonality of purpose*.

(13.24b) **That Red Riding Hood, Grandma, and the Wolf now felt a certain commonality of purpose* just so happened.

Cleft constructions

Clefting has the effect of ‘cleaving’ an original sentence into two clauses as a way to shift the focus of interest. There are many different types of clefts, but we will focus on one: ***it*-clefts**.

Take the sample sentence *Grandma, I have brought you some healthful snacks* (with constituents in the basic unmarked word order). But, let’s say Red Riding Hood wanted to highlight those healthful snacks. An *it*-cleft has the effect of splitting off this constituent and giving it prominence. Two clauses are formed, with the first introduced by an empty subject *it* and a form of *be*, followed by the focused constituent. The rest of the sentence is recast as a relative clause beginning with *that* (see Chapter 9 on relative clauses). It is an effective way of getting our attention by delaying mention of what is of special significance.

(13.25a)

[basic sentence] Grandma, I have brought you *some healthful snacks*.

(13.25b)

[it-cleft] Grandma, it’s *some healthful snacks* that I have brought you.

The advantage of using it-clefts is that they allow speakers and writers to show what is foregrounded and what is backgrounded. The backgrounded information is generally assumed to be known or presupposed. Don’t be surprised if these sound a little contrived – clefts require surrounding discourse to make sense. They might also have an Irish ring to them; Irish English is distinctive in showing a much wider range of clefting possibilities, including verb phrases and even complements – for example, *Is it tricking me you are* and *Because ‘tis a teetotaller he is*.

What types of presupposition can you relate to in the following examples?

(13.26a) [basic sentence] *Red Riding Hood set off with her basket through the woods.*
(13.26b) [foregrounding Red Riding Hood] *It was Red Riding Hood who set off with her basket through the woods.*
(13.26c) [foregrounding with her basket] *It was with her basket that Red Riding Hood set off through the woods.*
(13.26d) [foregrounding through the woods] *It was through the woods that Red Riding Hood set off with her basket.*

These might also sound odd because clefts are normally reduced. Old information requires minimal mention and one way of doing this is by omitting it altogether. So if we were to ask, *Who set off through the woods?*, the expected reply cleft would be, *It was Red Riding Hood*. The information in the relative clause (*who set off through the woods*) is given and doesn't need to be repeated.

Fronting

Fronting is a focus strategy that differs from the others discussed so far because it doesn't involve anything other than a straightforward word order change. All fronting does is move something to the front of the sentence; this is usually a topical element, or it forms some sort of contrast. In the following example, Red Riding Hood could have placed *your sexist remark* in initial position to give it special focus.

(13.27a)	[basic sentence]	I find your sexist remark offensive in the extreme.
(13.27b)	[fronted object]	*Your sexist remark* I find offensive in the extreme.

You have probably recognised fronting as the characteristic feature of Yoda-speak (see Chapter 9). In Standard English, we don't have quite the same creative potential as Yoda, yet lots of different material can occur up front using this strategy.

(13.28a)	[basic sentence]	They lived together in the woods happily ever after.
(13.28b)	[fronted adverbial]	*In the woods* they lived together happily ever after.

Fronting can also trigger a reversal of the normal word order so that the subject follows the verb. **Inversion** moves the subject out of its natural environment; it shifts the emphasis and therefore represents another aspect of this focus device. In earlier English, this inverted order had considerable dramatic force and was typical of lively narrative sequences. It has retained a kind of mock dramatic effect, which writers still exploit today.

(13.29a)	[basic sentence]	The wolf burst into the house.
(13.29b)	[fronted adverbial with inversion]	*Into the house* burst the wolf.
(13.30a)	[basic clause]	Red Riding Hood walked on along the main path.
(13.30b)	[fronted adverbial with inversion]	*Along the main path* walked Red Riding Hood.

Dislocation

Dislocation moves an element to the left- or right-hand end of the sentence, leaving behind some sort of copy (for example, a pronoun) in the gap left by the fronted constituent. Prosody or punctuation indicates that the element is not part of the main clause. The purpose of the dislocated element is to focus the topic or to provide clarification. In the following example of **left-dislocation**, the fronted constituent (*your sexist remark*) is kind of 'hanging' outside the sentence structure, and a pronoun copy (*it*) appears in the original position.

(13.31a) [basic sentence] I find your sexist remark offensive in the extreme.
(13.31b) [left-dislocation of object] *Your sexist remark*, I find *it* offensive in the extreme.

The grammatical relations have not changed; 'I' and 'your sexist remark' are still the subject and object respectively, but the object has become detached from the rest of the sentence, separated by a distinct pause (indicated here by the comma). So, it is different from, say, passives, which must change a constituent into a new subject in order to be able to front it. In writing, this construction is extremely rare, and it is heavily marked. Speech, however, is built much more along discourse lines and, as we will soon see, left-dislocation structures like these are common.

Right-dislocation involves the same process as left-dislocation, but with the rightwards movement of material; in other words, it goes to the end of the sentence rather than the beginning. This strategy is also used for clarification, or else for special focus.

(13.32a) [basic sentence] He put on Grandma's nightclothes.
(13.32b) [right-dislocation of object] He put *them* on, *Grandma's nightclothes*.

You might imagine someone saying here, 'He put them on (can you believe it) – Grandma's nightclothes!' Notice how the right-dislocated element is separated off from the rest of the clause by a distinct pause, often indicated by a comma (or dash). Like clefting, extraposition and *there*-constructions, right-dislocation introduces a pronoun that announces what is coming up in the discourse (but recall that these other pronouns aren't referential, just empty placeholders). It's a good way of building up expectations.

In conclusion, you can see that there are times when these various principles clash. The desire to have old information precede new is strong, and usually this decides the grammatical structure used to convey a particular message. So, there are strategies to manoeuvre new information into an end-of-sentence position. But, what if that new information is so important that it warrants special attention? To take something out of its customary position is always an effective way of putting focus on it and there are also strategies that do this job. Just as an aside, we acknowledge that it's sometimes not easy to distinguish between information that has been put in a special position because it is new and important, and information that has been put in a special position because it has been focused on. Think of focused information as emphasised information. The sentences that contain a structural focus position are quite different from the sentences in which we have a simple 'old precedes new' ordering.

13.3 Cohesion

At this stage, we move beyond thinking about information structure and associated structures and categories of grammar, and turn to consider other qualities that help readers or listeners to make sense of texts. **Cohesion** is concerned specifically with the linguistic relationships that hold within the text: how meaning is built up, and how connections are woven word by word and sentence by sentence. There are a number of ways of making sure connections between sentences are clear.

- **Anaphora** refers to relationships between words and other things that have already been said in the interaction (or written in the text). Mostly, when people talk about anaphora, they are thinking of things like pronouns – unless you're already clear about who *he, it, she* and *them* are in *Where did he leave* it? or *What has she done with them?*, it is impossible to know how the sentences should be understood. Here, anaphora refers to the link between the pronouns and a previous mention of that entity in the conversation or text. Anaphora can involve all sorts of expressions, not just pronouns. For example, in *I remember her from Perth but she doesn't work there anymore*, the anaphor *there* refers back to the phrase *in Perth* and is, therefore, a 'pro-prepositional phrase'. Noun phrases can also have anaphoric relationships with previous noun phrases (*I saw my grandfather the other day. The old fellow's really quite frail now*).
- **Ellipsis** refers to the way that words can be omitted if they have just been used and repeating them would be redundant. For example, in responding to the question *Where are you going now?*, the more natural answer is *To the movies* not *We are going to the movies now*. Although it may seem strange to

think that leaving something out can create a sense of connection, it helps to reflect that the omission indicates the semantic relationship is a very strong one.

- **Lexical cohesion** takes advantage of the many relationships among words that can be used to refer to the same thing. It turns repetition into a virtue by using repeated references to the same entity as an opportunity to introduce nuanced ways of describing it. For example, a text that refers to *the muddle-headed wombat, wombat, square animal*, and so on is both more engaging and more informative than one that that sticks to *wombat* all the way through. Since each time these words are used they refer back to the same character, they make links between different parts of the text at the same time as they extend the meanings being encoded.
- **Conjunction** refers to links between sentences made by words such as *and, because* and *while*.

Johnston and Schembri describe how these strategies are used in establishing and maintaining cohesion in Auslan. This includes the use of full noun phrases to establish referents, anaphoric devices such as pronouns, ellipses when the referent is well established, as well as strategies that make use of Auslan-specific resources such as modifications to the verb (for example, by signing in a particular space associated with an established referent).

Anaphora can be funny

Medicines can be harmful to young children. Make sure you keep them locked in the bathroom cabinet. When the previous context is available, anaphoric expressions are easy to resolve. But, sometimes there can be ambiguity. Lots of jokes rely on the strategy of assigning the wrong referent to an anaphoric expression. In the warning about medicines above, people tend to first think that it's young children that they're being told to lock in bathroom cabinets – mainly because that was the last noun phrase in the preceding sentence. Obviously, it's not a good idea to do that! There are lots more corny jokes out there that play on anaphora. See if you can work out what's going on the following three examples.

(1) *Doctor:* How are you feeling today?
Patient: My breathing still troubles me.

Doctor: Mm! We must put a stop to that.

(2) *Arts 2004* will feature an exhibition of over 1200 paintings by young artists. They were all executed over the past two years.

(3) Fried eggs should be cooked properly and if there are frail or elderly people in the house, they should be hard-boiled.

Reflexive pronouns

Reflexive pronouns make up a special category within anaphora, because they specifically make the link between antecedents and themselves. Think of them as reflecting (or being 'reflexes' of) noun phrases that occur elsewhere in the same clause. Compare example 13.33 with example 13.34.

(13.33) Sarah bumped into Jodi and knocked herself out.

(13.34) Sarah bumped into Jodi and knocked her out.

In English, reflexive pronouns can also be used to add emphasis. We will use another overworked joke to illustrate: *Tired of cleaning yourself? Let our trained staff do it for you*. The ambiguity here occurs because *yourself* could be understood reflexively (*you clean yourself*) or emphatically (*you yourself clean*).

Different languages have different strategies for dealing with reflexive relationships between the pronoun and the preceding noun phrase. Here are a few from our case-study languages.

Gurindji

Gurindji makes use of one reflexive pronoun with the form *-nyunu*. It's a suffix that can either occur on the verb or in a group with an independent subject pronoun.

(13.35)

Ngurnpurlkanyunu
1exclusive.plural-reflexive treat.this.one
‘We treat ourselves with this one.’

You can see what a non-reflexive sentence using the same verb looks like in example 13.36. In this example, the pronoun *ngurnayinangulu* encodes both the first-person plural exclusive subject (*we*, not *you*) and the third-person plural object (*them*).

(13.36)

Ngurnayinangulu
1exclusive.plural.subject.3plural.object treat child
‘We treat the kids.’

Japanese

In Japanese, there is a single reflexive form that is used, regardless of the person and number of the referent. As with English reflexive pronouns, there is often an emphatic overtone (compare with *Do it yourself* with *He only cares about himself*).

(13.37)

kangae-nasai
reflexive think.about-polite.imperative
‘Figure it out yourself.’

(13.38)

kawaia
person
‘That person has a big head.’

(Literally: That person thinks only themselves are cute.)

Mali

In Mali, it is more useful to think about reflexive meaning as being assigned to the verb as a derivational suffix, rather than to a pronoun. Compare the non-reflexive in example 13.39 with the reflexive in example 13.40.

(13.39)

ka hāvēt-ki
relative look.after-she
‘who looks after it,’ (referring to a noun in the feminine gender)

(13.40)

[illegible]
[illegible]
‘My wife, you must look after yourself properly.’

Note that the pronoun and the derivational suffix occur in the same slot after the verb. One reason why *-nas* is not treated as a reflexive pronoun (as in Gurindji or Japanese) is that for verbs of movement it is necessary to use an additional suffix *-na* that reduces the transitivity of the stem: *suchut* ‘push, shove’, *suchutnanas* ‘ease [push self] through’. This indicates that there is something more than pronoun replacement going on.

Deixis: ‘Pointing words’

When we talk to each other, we rely quite a lot on knowing where we are, what time it is and what has been happening around us. This involves something called **deixis**, the relationship between the context and the way we interpret words such as *here, there, now, then, me* and *you*. The term **deictic** simply means ‘pointing’. Normally, deixis applies to spoken language, but it can occur outside of speech (for example, in dialogues from written texts).

Consider the following sentence:

(13.41) Yesterday I saw you here with those things you like so much.

This example has a number of words that only make sense if the listener knows the circumstances of the utterance.

- Which day is *today* and hence which day was *yesterday*?
- Who is speaking (*me*)?
- Who is listening (*you*)?
- Where is *here*?
- What things are *those*?

A useful concept for organising all the things that you need to know in order to understand deictic forms is the notion of the **deictic centre**. Usually the speaker, and the time and place of speech are in the centre (for example, *I, now here, this*) and other related forms are more distant (for instance, *you, they, then, earlier, later, there, elsewhere, that, something else*). But, in storytelling or in speech exchanges, the deictic centre can shift to other participants.

Deixis can be funny too …

One of the interesting things about anaphora and deixis is that some forms in English can fill either anaphoric or deictic functions. In the following joke, the word *this* is used deictically by the man making the phone call and anaphorically by the nurse (with reference to the baby implied as being expected by use of the word *labour*). Both uses are acceptable and the ambiguity creates room for humour.

> A man phones the local hospital and says urgently, ‘You’ve gotta help, my wife’s in labour!’ The nurse says, ‘Calm down. Is this her first child?’ He replies, ‘No! This is her husband!’

Here’s another joke along similar lines:

> A man walks into a book shop and says, ‘Can I have a book by Shakespeare?’ ‘Of course, sir’, says the salesman. ‘Which one?’ The man replies, ‘William’.

In this case, the word *one* is ambiguous. It might refer anaphorically to (a series of different people called) Shakespeare, which is how it is understood by the customer. However, this use is not the most likely – deictically it refers to the books by Shakespeare, some of which may just be in the shop.

Another feature of cohesive texts (those that fit together linguistically) is that they have to **cohere** (fit together semantically). This can be a challenge. We are sure you've had similar experiences to this next example:

(13.42) [*Doorbell rings*.]

Fritz: That's the door!

Siggy: I'm in the shower!

Fritz: Okay.

We'll return to this example in Chapter 14 and consider how it makes sense in context even though it does not appear that Fritz and Siggy are even in the same conversation.

13.4 Discourse analysis

There is a large and diverse body of work in linguistics and related fields concerned with the way meaning is made through language. Discourse analysis is the umbrella term for all of this research. Although discourse analysis includes a concern with the structure of larger stretches of text (and talk), it also encompasses thinking about the ways in which social order is produced through interactions between texts (and talk) and other aspects of social life.

In this chapter, we focus on some of the more structurally oriented topics associated with discourse analysis before turning to more ideologically oriented understandings of discourse analysis in Chapter 15.

Big D and little d discourse analysis

> Discourse is a 'dance' that exists in the abstract as a coordinated pattern of words, deeds, values, beliefs, symbols, tools, objects, times, and places in the here and now as a performance that is recognizable as just such a

> coordination. Like a dance, the performance here and now is never exactly the same. It all comes down, often, to what the 'masters of the dance' will allow to be recognised or will be forced to recognize as a possible instantiation of the dance.
>
> Gee (2014), *An Introduction to Discourse Analysis: Theory and Method,* p. 53

Many researchers across the social sciences and beyond have embraced the concept of Discourse analysis (with a big D) as it relates to these larger concerns with the production of social order. Before we turn to think here in more detail about the structural organisation of discourse (that is, discourse with a small d), let's take a moment to consider the connections between the two.

In recent work, Tricia Cleland Silva examined the discourse around the transnational labour market of nurses, specifically the business of recruiting Filipino nurses. By analysing interviews and documents, she looked at how organisations and their representatives packaged Filipino nurses as marketable products. Consider the following statement by a representative from a private recruitment company that sources Filipino nurses for employment in Finland:

> But, the Philippines for us is, I suppose, in some ways, like it is a natural, I don't want to use the word resource – it sounds terrible – but, for us, we would go to the Philippines because one, the training is good there. We know that we are not going to have to search too hard for the nurses there and they are-they have this ability to sell in other countries very quickly, just like the saying that you would be aware of, that Tagalog saying, 'there is no bread too hard for warm coffee'. And that is a wonderful saying from the Filipino point of view that they will go anywhere. Just treat them right, look after them, and they will pay you back tenfold.

The commodification of the labour provided by Filipino workers ('nurse packaging') is hard to miss in this quote: they are presented as a natural resource that can be sold on the international market with an outstanding return on investment (acknowledging that a thing is terrible to say is not the same as choosing not to say it!). In thinking about the way these meanings about Filipino migrant nurses in Finland are made, we focus on the ways in which they are described, and the connections and feelings the description evokes – these are concerns associated with big D discourse.

One of the funny things about the contrast between big D discourse and small d discourse is that no one is very active in pursuing the label small d discourse. (Given the connotations of relative importance associated with that capital D, perhaps this is not really so surprising.) Rather than embracing this label ourselves, or asserting that it belongs to any particular approach outlined below, we will stick with our strategy of focusing here on the structural aspects of discourse, leaving the more ideological aspects to Chapter 15.

Discourse markers

One important set of words that feature prominently in discussions of structurally oriented discourse analysis comprises discourse markers. These little words can fall into a wide range of grammatical niches, but share the function of making connections for the listener between what has been happening in different parts of the text or talk.

In a lovely study about *oh,* Schiffrin shows that this particle has a range of functions associated with acknowledging and navigating the request for and receipt of information; she argues that it appears to buy individuals time to focus temporarily on problem-solving informational issues while still acknowledging interactional and social needs. One of the newer discourse markers on the block is *yeah-no.* Burridge and Florey show that it has an important linking or orientation role in our conversations, acknowledging what has preceded and resuming components of the conversation from earlier in the discourse. By marking the connection, it therefore provides coherence. For example, we might be chatting about discourse markers generally, then suddenly someone says: ‘Yeah-no, Silva’s work really illuminates how those nurses were socially constructed for their Finnish clients.’

If you’re learning a second language, getting a grip on discourse markers is a great way to sound like you know what you are doing! They are notoriously difficult to define or translate, but here are some rough definitions of a few discourse markers from Japanese to give you a sense of what it’s all about. (These markers are complex creatures and we can’t do justice to this complexity here; we urge you to seek out the references given in further reading at the end of the chapter.)

- *ne* indicating a request for confirmation or agreement
- *yo* indicating certainty
- *na* indicating emotion (used informally by men)
- *sa* indicating emphasis (used informally by men)
- *wa* indicating emphasis (used by female speakers).

In addition to establishing discourse connections, you can see from this handful of Japanese examples that these devices also have functions that relate more to the speaker and can be more personal. It is true that people often sneer at discourse markers, mistaking them for meaningless little expressions that speakers use to fill in time while they are thinking. We hope you can see that they are anything but. Discourse markers are vital to interpersonal communication, whether capturing connections between pieces of discourse or connections between the speakers themselves.

Discourse markers in science writing

Because discourse markers are used to structure discourse, you would probably expect that some of their patterns of use would be carried over into second-language texts. To test this hypothesis, Rezanova and Kogut compared a small corpus of scientific texts about geology written in both Russian and German. They found that whereas Russian scientific discourse seeks to reassure the reader using expressions that can be translated with English phrases such as *it goes without saying, beyond doubt, undoubtedly, it is safe to say*, the authors of the German publications instead seek to hedge their assertions, with words equivalent to English *probably, possibly, it can be assumed*. The authors of the research suggest that this different way of positioning the writer in relation to the ideas in the papers may reflect the need of the German authors to be moderate in the face of a strong tradition of objective critical commentary in Western European scientific tradition compared with a more discrete approach to criticism in the Russian tradition. We again see the ways in which local text-level language practices are readily connected to broader social practices.

13.5 Different types of communication

> No one means all he says, and yet very few say all they mean, for words are slippery and thought is viscous.
>
> Henry Adams (1907), 'The Grammar of Science', in *The Education of Henry Adams*

In the preliminary parts of this chapter, we have referred to two different modes of communicating, namely speaking and writing, as well as making some reference to the distinct mode of signing. In this section, we consider similarities and differences between speaking and writing and the ways in which these differences impact other aspects of communication.

There are all sorts of differences (linguistic, functional and social) between speech and writing, and we have summarised some of them in Table 13.2. Not all examples of speech or writing will have these features, though, and we address this issue a little later. In particular, as you work through the table, think about how the digital revolution has blurred the traditional distinction between speech and writing. Online chat is written, of course, but it shares many of the features, not just of spoken language, but of casual conversation. In real-time online communication, people exchange messages in much the same way as they would do when chatting face to face.

Table 13.2 Some differences between speech and writing

Speech	**Writing**
Lower prestige in society	Higher prestige in society
Social	Solitary
Dynamic and transient	Static and permanent
Typically more informal and intimate	Typically planned and edited
Typically face-to-face interaction	No visual contact, no contextual support
Immediate feedback	No immediate feedback
Basic units are sounds	Basic units are letters (alphabetic scripts) or symbols (syllabic or logographic scripts)
Supported by prosodic features (e.g. intonation, pitch, speed); also non-verbal gestures, eye contact	Few prosodic features captured by graphic conventions (e.g. punctuation, typographic forms, handwriting)
Pauses and silence	Gaps and dashes
Topic changes marked by discourse expressions (e.g. *yeah, so, right, well, now*)	Topic changes marked by headings, new chapters, sections, capitalisation, and so on; also expressions such as *first, finally, in conclusion*
Filled pauses such as *er, erm*	Thinking time and hesitations not generally indicated
Monitoring checks on listener involvement and maintenance of interaction (e.g. tags such as *isn't it*, rising intonation or expressions like *you know*?)	Occasional checks on reader attention (e.g. expressions such as *recall, try to bear in mind*)

Consider the examples in Table 13.3. It arranges various sorts of communication in relation to some of these parameters. Note that the example of the written mode in each set is represented in the shaded line, and the spoken is below it (unshaded). There are many more text types you could add here, but the relationship between them is more complicated than this table suggests – usually it involves a number of parameters.

Table 13.3 Classifying language in use

1 *Durability*:	
Permanent	**Ephemeral**
Published books	[illegible] notes
Recordings	Elevator small talk
2 *Formality*:	

[illegible] Legal contracts

Political speeches Chit-chat

3 *Standardisation*:

[illegible]

News broadcasts Nursery speak

4 *Interaction*:

[illegible]

Broadcasts Verbal duelling

5.*Function*:

[illegible]

Friendly banter

6 *Situation*:

[illegible]

Storytelling

7 *Spontaneity*:

[illegible]

Lectures Spoken conversations

8 *Privacy*:

[illegible] letters

Sermons Pillow talk

You'll see that we have included some examples of e-communication here. Emails, tweets, blogs, SMS texts, chatrooms and so on are now routine aspects of many people's lives, and are good examples of the blurred nature of speech and writing. These types of text involve writing, but they are interactive and therefore share many of the features of speech. The spontaneity and speed of this sort of communication mean that we simply don't go in for the same careful organisation and planning as we would with normal writing. Very little editing goes on. Talk becomes written and this means a much looser construction, with repetition, false starts, digressions, comment clauses and asides. As we discussed in Chapter 10, electronic communication of this kind also has the expressive opportunities in word order that we typically associate with spoken language (for example, *I get called a lot of oder names coz my name is kinda hard to pronounce so yeah I dont mind wat ever*), and there is slang and grammatical informality as well (for example, *bestest everest friend*). Special graphic devices, such as emoticons and emojis, communicate something of the same meaning conveyed by the prosodic and paralinguistic features of speech, and make up for the fact that we tend not to give the language the careful wording we might in more traditional types of writing. Typically, people bang down their ideas with the speed and spontaneity of speech, but without the full support of the expressive devices that speech can utilise. And, there are rarely the same time-consuming politeness routines that are also important when it comes to establishing social rapport (greetings, leave-takings, and so on). Without this phatic lubricant to oil the interaction, messages can come across more brusquely and more directly than originally intended.

Having established just how different the possibilities and limitations of speech and writing may be (even though they have so much communicative function in common), we turn now to look in more detail at two specific types of communication: conversations and narratives. We can only scrape the surface here, but in the Further reading section at the end of this chapter, you will find mention of recent research that has identified more refined sets of conversational interactions and narrative types, each with a distinctive set of categories.

13.6 Conversations

As we noted above, casual conversation is feted as the prototype of spontaneous speech. In its ideal form, it involves unplanned interaction between socially equal participants, in an informal setting. Some researchers have questioned the importance given to casual conversation, noting that for many people this is not, in fact, the type of language they use most often. Stage of life, occupation and other related social categorisations have a great deal to do with the kinds of communication that are most frequent for particular individuals.

Nevertheless, an important characteristic of casual conversation is its divergence in many important ways from prototypical written language, which was the default for language analysis for some time. Indeed, until quite recently, relatively little was known about the features of spontaneous chat; the activities of linguists and grammarians overwhelmingly concentrated on the structures of planned and highly standardised (and therefore usually written) language, and virtually ignored the unplanned and more spontaneous discourse varieties. One of the reasons for this was practical. Recording is a modern phenomenon and before it became widely available, we simply did not have proper access to reliable, reusable corpora of live unsolicited speech. Unfortunately, the legacy of this is that people have come to equate ‘normal’ language with formal (usually written) language. This, together with the long-established importance of writing over speech in society, has meant that informal spoken features, such as you’ll see in a moment, remain highly stigmatised.

An obvious challenge in studying spontaneous speech is that casual conversation is transitory and impermanent. Doing analysis on a casual conversation and then reporting findings to others requires that the researcher must find a way to fix the interaction, and make it visible and permanent. Writing it down is the obvious response – but the challenge is to make a written record that leaves in all the information that writing normally excludes (pauses, coughs, sniffs, marked pronunciations, changes in volume and so on). Researchers have developed transcription systems – ways of representing speech – that preserve at least some of its 'performative' aspects. There are several different systems around for use in this task – one is presented below. Notice that this system uses punctuation marks to signal intonation boundaries, stress and so on rather than grammatical boundaries. Table 13.4 provides a summary of the ways in which qualities of speech are rendered in this system.

Table 13.4 Transcription symbols

.	Final falling/level intonation unit
,	Continuing intonation unit
?	Final rising intonation unit (on a question)
¿	Rising intonation (on a declarative)
word WORD	Underscoring indicates loud; capitals very loud
(.)	Micro-pause within or between utterances
(0.0)	Time elapsed by tenths of seconds
wor-	Truncated word
wo:rd	Lengthened word
wo(h)rd	Laughter within word
.hh; hh	Inbreath and outbreath
(word)	Uncertain transcription
( )	Unclear/indecipherable talk
↑word; ↓**word**	Onset of noticeable pitch rise or fall
°word°	Material between degree signs is quiet
<words>; >words<	Outwards arrows for faster speech; inward arrows for slower speech
word=; =word	No discernible pause between turns
[ ]	Overlapping speech
@	Pulse of laughter

The following transcription is an extract from a story told by a women's basketball player to her team-mates during a warm-up session. You can hear the recording from which it was made in the online materials for this book. It is important to bear in mind that, in practical terms, a transcription like this can never capture everything about a spoken interaction. Researchers always make choices about transcription conventions (usually based on their interests). Transcription is a labour-intensive process.

262. corb: AND THE FIhhRST THING HE DI:D, like- I stayed
263. there, on Sunday night at the hospital,
264. (0.8)
265. corb: u[m–

266. tan: [with him?
267. corb: two chairs/
268. tan: yeh/
269. corb: and he wo- >the first thing he did when he woke up
270. out of the anaesthetic< was ºbhhecº! (0.4) %f =
271. corb: = be:c! [((chuckles))
272. jb? [º o : : =
273. jb? = [: : : h .º
274. tan? º[that's so cu[te.º
275. lang: >[and you went< º=
276. lang: = [yesº! (and sorta walked out going YE:::HH) ((clapping))
277. jb? [o : : : : h,
278. corb: his mum's going – Shane, your mum's her[e/
279. tan:[yAh @ @ @
280. corb: ((laughter))
281. tan:((laughter))
282. ? ((clapping))
283. corb: .HH[:: I'M LIKE – THAT'S ALRIGHT, I'm just going =
284. jb: [YEH LIKE WELL (I'M IN HERE/)
285. corb: = to have a seat, my job's [here done. ((note
286. weird word order!!))
287. jb: [ye@h!
288. (2.8)
289. sev: ((laughter))
290. tan?: hu::[:nh.
291. ? ↑[o h w [: /
292. corb? [((sniff))
293. corb: so yeh, it was good.

Following are a few important features of casual conversation that are apparent in the transcript.

Expressing support

Conversations are collaborative efforts, as our excerpt here certainly shows. As participants alternate between the roles of speaker and hearer, they work in partnership. Hearers collaborate in extending the turn by not taking over the floor. They also contribute to the story by indicating they are listening and by egging on the speaker with facial expressions, nods and smiles, laughter and encouraging noises such as *hmm, yeah, right* and *ooh*, and so on. These supportive cooperative noises are sometimes called minimal responses or **back-channelling**. In the above excerpt, much of the overlapping speech is also intended to express support for the main speaker.

Turn-taking

One of the delicate things about having a chat with someone is figuring out when to start talking and when to stop doing so. We have to monitor our conversational partners and try to work out when a change of speaker is possible. If there's a breakdown, we also have to handle the problem and somehow repair it. All this can be difficult (especially when people don't know each other well). Sometimes, there are gaps and often more than one person can be heard speaking at a time. Conversations have occasionally been described as 'two voices collaborating as one' (for example, speaking all at once or over the top of each other can express enthusiastic sympathy, as in our excerpt), but at other times it can be a strong-arm tactic for gaining the floor.

A useful conceptual tool for thinking about turns is the **transition relevance place**. This refers to points in a conversation where it makes sense for there to be a change in the speaker. Line 271 in our excerpt is a good example of this. The speaker has made her point (that after waking from a general anaesthetic, her boyfriend called out for her rather than for his mother) and her chuckle is a signal that she has finished speaking, leaving the floor open for others to respond. Other types of signals that a transition relevance place has been reached include prosodic markers (such as falling intonation or the drawing out of a word's final syllable) or non-verbal cues (for example, the speaker's patterns of breaking or maintaining gaze, turning away, leaning back).

A lot of talk involves sequences of what are called **adjacency pairs**. These are adjacent turns that have a close relationship with each other. They help with the general ebb and flow of a conversation, by signalling openings and closings, for example, or marking moves within a conversation. Take the automatic patterns you find when people greet each other or say goodbye. It is actually quite hard to initiate or close a conversation without this sort of routine – this is why you find that talkback callers on radio always ask presenters how they are before they make their comment. As in this greeting–greeting exchange, these sequences usually come in pairs, with one utterance triggering a response: question and answer; compliment and acknowledgement; thank you and response; apology and acceptance; complaint and excuse. Our previous excerpt includes a question–response pair in lines 266 and 267, where a member of the audience of the story asks for more detail about where the main speaker had slept.

The official record

Many transcriptions kept as a matter of public record omit a great deal that is actually characteristic of spontaneous speech. Consider the following example from the New Zealand Parliament *Hansard* (Question Time, 28 October 2014).

> **Rt Hon WP:** Consistent with an open and transparent Government, when will the Prime Minister give the public all the facts about his staff's services to the National Party's Jo de Joux whilst employed at the public's expense?
>
> **Rt Hon JK:** I have not seen any advice that would confirm that the National Party has done anything other than spend taxpayers' money appropriately and legally.
>
> **Hon Dr NS:** What did Brendan Horan say? Look at your own house.
>
> **Rt Hon WP:** Is it not a fact – well, actually, we were exonerated, turkey.
>
> **Mr Speaker:** Order! [*Interruption*] Order! That is a good lesson to that quarter of the House of what happens when you get an interjection through a question. Would the right honourable gentleman simply ask his supplementary question?
>
> **Rt Hon WP:** Is Jo de Joux, a National Party campaign manager, deeply implicated in improperly acquiring services from his office's staff?
>
> **Mr Speaker:** I invite the Prime Minister, if he wishes, to answer, but I cannot see any prime ministerial responsibility in that question. I will leave it for the Prime Minister to answer it.
>
> **Rt Hon JK:** I have no authority for that, and I refute the proposition.

Which characteristics of spontaneous speech are missing here? How many of them do you think were present when the people involved were actually speaking? Why?

13.7 Narrative structures

Just as we can examine conversations and find patterns that people use to organise interaction, it is also possible to identify patterns in the ways we tell stories – both spoken and written. Labov and Waletzky traced out the typical features of narratives, using a series of hundreds of interviews in which English speakers were asked to describe a significant event from their lives. On this basis, they identified six stages as basic to narratives (see Table 13.5).

Table 13.5 Narrative structure

Narrative category	**Narrative question**	**Function and/or content**
Abstract (optional)	What was this about?	Engage the attention of the listener
Orientation	Who or what is involved in the story, and when and where did it take place?	Provide basic information to the listener
Complicating action	Then what happened?	Describe the events involved in the story
Resolution	What finally happened?	Reveal the final key event
Evaluation	So what?	Make the point of the story clear
Coda (optional)	How does it all end?	Mark the completion of the story

The extent to which it is reasonable or fair to take this model of narrative and apply it to speakers of other languages, or even to speakers of English from other cultural backgrounds, remains unclear. Although some research acknowledges that narratives may involve ‘characters engaged in goal-directed actions that relate to and/or resolve problems or complications’, it is also true that there is a good deal of variation in ‘the nature of the problems that get resolved, the nature of the relationships between the characters, and the types of narrator explanations’. Others have noted that this model is strongly event-focused, and fails to account for narratives that report experiences. One reason why there is an event focus in this particular model is that it is based on narratives about incidents, as we mentioned above.

Subsequent researchers interested in conversational interaction have extended the model by Labov and Waletzky to encompass a wider range of interaction types. One strategy for expanding the range of structures that we know about is to explore texts based on a variety of features, including spoken or written language, or material from different contexts, intended for different purposes. In Chapter 10, we described linguistic differences associated with different social settings in terms of register – the types of grammar and word choices typically involved. There are also differences in terms of text structure that can be useful to recognise and explore. Recipes, for example, have a particular structure that can be expressed in similar terms to the narrative structure outlined above (see Table 13.6).

Table 13.6 Recipe structure

Recipe category	**Cook's question**	**Function and/or content**
Title	What will I cook?	Invite the reader to consider making the recipe
Photo (optional)	What's it supposed to look like?	Provide the reader with more information about how the recipe will look

		when finished; provides a further invitation to the reader
Description (optional)	Why should I cook it?	Describes the dish, often with reference to a personal experience of the recipe writer
Ingredients	What's in it?	Lists ingredients and required amounts (may include preliminary preparation notes)
Method	How do I cook it?	Step-by-step description of how dish is made (sometimes numbered, sometimes with photographs of steps)
Serving suggestion (optional)	Anything else I need to know?	Additional ideas about how to serve the recipe
Dietary information (optional)	Can I serve it to my gluten-free/dairy-free guests? Will it make me fat?	Information about the type of ingredients, number of calories and so on

Cultural aspects of narratives

In our experience doing fieldwork in different cultural settings, we have observed that the types of narratives we value are not the same as those valued in other places. For example, in the Mali community the location of the narrated events (that is, place) is much more important in providing orientation to the listener than time – a strategy that Tonya found quite confusing during her fieldwork, partly because much of the country around Marunga was not well known to her, but especially because her expectation was that she would be able to use time to organise how she understood the narrative. She often wondered: Is this a story from World War II? Is this a story from when the speaker was a child? and Is this something that happened recently? This last question was especially difficult to put aside since asking about whether something was recent rarely prompted a helpful response, presumably because it is a highly relative concept. Without this temporal orientation, Tonya often experienced these stories as unanchored and confusing (a reflection on her competence as a listener based on her status as a cultural outsider, rather than a reflection on the skills of those who so generously agreed to work with her!).

One strategy for comparing narratives across cultures is to use picture stories without words to prompt narratives from participants. Researchers such as Saville-Troike report variation in the following areas of narratives obtained in this way: openings and closings; the treatment of characters, how they are related to each other and what thoughts and feelings are attributed to them; omission of some expected events and addition of other events, not expected; additional comment based on background knowledge; symbolic interpretation of story elements; and evaluations of characters' behaviour. This research confirms that narrative structures vary across cultures and that skilful communication is not the same everywhere.

Narrative structure as normal

Notwithstanding the complications of constructing appropriate narratives in specific cultural settings, the Narrative Scoring Scheme (NSS) is used in many English-speaking contexts, including Australia, to evaluate the linguistic development of children by examining their performance as narrators. It invites the tester to assign children's performance into one of three categories (proficient, emerging and minimal/immature) based on the extent to which they produce narrative in expected ways. The test does not make reference to different types of storytelling, but makes specific assumptions about what effective narration looks like. These include the expectation that the child will provide temporal orientation, narrate some of the story in the character's voice, provide smooth transitions between events, use a variety of mental-state verbs (that is, report what characters thought and felt) and use a general concluding statement at the end of the story. (It is noteworthy that this final requirement has been treated as optional in most theoretical writing on narrative structure, but is a required category in this assessment tool.)

One study of Aboriginal children in a North Queensland school used this scoring mechanism. After noting that Australian children in general are disadvantaged by standardised tests since so few of them have been normed for the Australian population, they noted that the Aboriginal children they tested generally:

> did well with the Introduction, Character development and Conclusion components. This is consistent with the observation that people and place are of key importance in Aboriginal storytelling … It is unclear why most participants had difficulty describing the mental states of the story characters (feelings and thoughts). Referencing, Cohesion, and Conflict resolution were challenging components for the younger children but older children had good mastery of these. This suggests that experience with narratives at school contributed to the ability of Aboriginal children to produce more mature oral narratives as measured by the NSS.

Recognition of the ways in which culture influences what is accessible, relevant or exciting in a narrative may lead to greater understanding of these differences. Alternative perspectives include the idea that older children have had more opportunities to learn to perform the types of narratives expected in non-Aboriginal settings such as school, and this facilitates their more culturally compliant performance on the NSS rather than being an indication of the maturity of their narratives *per se*. This is a confusion built into the language of the assessment tool, which rates non-conforming narratives as ‘immature’. A more carefully constructed interpretation of results would provide researchers with the opportunity to disentangle acculturation from maturation in a nuanced way. In addition, better knowledge of the narrative practices that children bring to school would support teachers and other professionals in making fairer evaluations of narrative ability, and allow children to be acknowledged for the full spectrum of their developing cultural competence in both spheres.

In addition to structure, languages also show differences in the linguistic aspects of their narratives. German, for example, has a particular form of the past tense (the imperfect) that is strongly associated with writing and another (the perfect) that is more often associated with talk. Learning to use both types of verbs can pose a second-language learner of German with some challenges. Depending on the situation in which learning happens, it is easy to end up more practised in one set of forms than the other.

Mali narratives are characterised by something called ‘tail-head linkage’. In the following example, there are two stretches of talk linked by stretches of repetition. The first stretch describes the journey of some spies into the bush in lines (a–c). The second stretch describes the capture of some people in the bush in lines (d–f). The repetition at line (d) serves to link the two events. An expansion of line (d) is given in line (e) and this is also managed using repetition. (Because the focus here is on the development of the narrative rather than particular words or phrases, glosses have not been included.)

(13.43)

(a) *a-chases-ta avathē vik ta met na-pēt kama a-is-ka i-na-mēk*

‘Those particular spies they went up from the road.’

(b) *da ta thon.dēm ma Mēndai sa-i-vuk*

‘and they went along the Mundai River up’

(c) *da ta muēn pēt kama a-bang-ēm ma ta tes ngam ia Maslabe*

‘and they came to the village which they call Maslabe.’

(d) *ta muēn.vēt ngam da ta chēthep kama a-ru-ta*

‘They found it [the village] and they captured the people.’

(e) *ta muēn.vēt ngam da ta chēthep kama ru-ta ama adēvaung na-ta*

‘They found it and they captured three people.’

(f) *da a-ru-ta ta mo da ta vang*

‘And all the other people ran away.’

This example also shows the tendency common to many Papuan languages of establishing participants early on and then minimising the use of noun phrases in the discussion of events. After the first sentence, all the clauses in this passage are limited to one noun phrase referent – typically an object or a prepositional phrase indicating location.

The frog story in Russian

One study that explored differences in narrative strategy across cultures invited a series of mother–child pairs to read together a book called *Frog, Where are You?* by Mercer Mayer. This book is used widely in psycholinguistics and field linguistics, and involves a story told in pictures rather than words. As described earlier, in these types of activities participants are invited to ‘read’ the book and are, therefore, free to make up the text in words as they go along. Researchers asked the following types of families to read:

- native speakers of English in the United States
- native speakers of Russian in the United States
- native speakers of Russian in Ukraine
- native speakers of Russian in Israel
- Hebrew-speaking Israelis living in Israel.

Their goal was to explore several aspects of the storytelling process, including the mothers’ story-telling style, story length, the amount of talk about emotions, and the frequency and types of questions to children related to their cultural background and life experiences, particularly the experiences of Russian-speaking immigrants in the United States and Israel. They anticipated that the Russian-speaking immigrants would show differences reflecting the cultural practices of the places where they had settled – that the people who had settled in Israel would tackle the story differently from the people who had settled in the United States.

One interesting aspect of the findings was that, while there were significant differences among the mothers from different cultures in terms of how they told the story, the children across the sample showed much less variation in their communicative choices. In relation to differences between the mothers from different groups, the authors report, for example, a higher number of emotion words in the narratives of Russian speakers; this supports the results of previous research suggesting that individuals of Russian language background focused more on emotions than the American mothers.

We return to this theme of how people use language to reflect and position themselves in relation to culture and identity in more detail in Chapter 14.

It is interesting to note that this narrative strategy is recognised within the Mali community as being culturally specific and failing to match the norms of literary English. In preparing a collection of stories for the community, Mali people removed instances of tail-head linkage from the texts that had originally been recorded as spoken performances. Although speakers wanted to use these structures in traditional spoken modes, the transition to writing seemed to also trigger a transition to a different set of narrative conventions in which the cohesive functions of tail-head linkage were no longer valued.

In this section, we have shown that there are relationships between context, culture, language and the possibilities for narrative structure. Narrative structure is not a universal phenomenon in anything but the very broadest sense, and narrative practices are not isolated from the cultural settings in which stories are told. Understanding how people use narrative, the stories they tell and the meanings they make, is more usefully considered as part of a wider set of cultural practices deployed in order to establish their place in the world and make sense of it. These cultural practices are also important in socialisation, and narrative provides a fertile source of linguistic modelling and cultural explanation.

13.8 Ethnography of communication

Our discussion of narrative has shown that there are many other factors besides the purpose of the communication that inform the ways in which people communicate with each other. We have also emphasised that these differences are related to culture. What is normal, expected, admired or abhorred is very much bounded by cultural expectations. By stepping back from a focus on a particular text type, we can begin to consider the social setting and the purposes for which communication is being used. Research in this area is known as **ethnography of communication**. (Ethnography refers to the study of people and culture; it is related to the more familiar word 'ethnic', which refers to people who identify themselves as belonging together for cultural or historic reasons.)

This method of analysing communication was originally called 'ethnography of speaking'; its creator, Dell Hymes (1972), came up with the mnemonic acronym SPEAKING as a framework to analyse the events of spoken communication ('speech events') in their cultural context as fully as possible. In this way, we can go beyond the content of the talk to consider how it relates to people and events surrounding it.

- **S**etting: describes when and where the speech act occurs, and extends to include information about the scene, defined as the 'psychological setting' or 'cultural definition' of a speech event.
- **P**articipants: the people present at the time of the speech event, including people who are addressed by the speaker as well as anyone else present; in some types of speech event (for example, a wedding ceremony) the roles may vary.
- **E**nds: involves the purposes, goals and outcomes of the event; these are related to why the speech event is happening and how it is being carried out.
- **A**ct sequence: refers to a more detailed account of how the event proceeds; relevant information here includes what is involved in completing the event, who speaks and when.
- **K**ey: includes the tone, manner or spirit of the speech event (for example, is the event light-hearted or serious?); these differences are expressed through the language that is chosen (including word choice, linguistic style, intonation and rate of speech).
- **I**nstrumentalities: what modality is used (speaking, signing, writing and so on)? What types of language (for example,

styles or dialects) are used? And how are they used?

- **N**orms of interaction: associated with the event and describe what expectations participants have of each other, what specific roles involve, how each person should behave and so on.
- **G**enre: is this a known type of speech event? For example, a wedding is a known type of event, with particular expected features.

See how many of these aspects of communication you can identify in the discussion of the role of English in Japan below.

Reconceptualising English as an international language

Increasing migration into Japan provides an opportunity to test the idea that English is the default international language in settings where people otherwise have no language in common. Ryuko Kubota and Sandra McKay conducted a project to evaluate the role of English in Japan in multilingual contexts. They chose, as their research site, a mid-sized city that hosts foreign workers from Brazil, China, Peru, Korea and Thailand. Many of these workers – particularly those from South America – are of Japanese descent. They quote one participant who comments on the lack of need for English translators within the foreign worker community at a school-based event:

> I was asked to attend to the English table as an interpreter. Approximately 50 people (students and parents) came. There were three Portuguese tables, a large Chinese table, a Spanish table, and a Thai table. None came to the English table.

The researchers found that, in practice, English is not the language of communication between local Japanese people and foreign workers – not least because foreign workers generally have not had access to opportunities to learn English. They argue that English is not neutral, but carries associations of race, class and national identity, such that the acquisition of English reinforces differences between the local Japanese population and foreign workers, rather than facilitating their removal. They conclude that the narrow focus on the acquisition of English in Japan has come at the expense of interest in other languages, and creates a situation of double monolingualism in Japanese and

English. They stress the importance of investigating the impact of the global spread of English, not only in terms of its use as a contact language or its threat to local languages, but also in terms of its relationship to other languages and cultures – how, for example, might the learning of English maintain or transform unequal racial, linguistic and cultural relations of power?

Ensuring that our descriptions of language behaviour are grounded in the real world requires an understanding of what people actually do (as opposed to what we think they do or guess they do), and a nuanced understanding of when, why and how people do things. Ethnographic approaches provide one framework for tackling this type of research.

Further reading

See Johnston and Schembri (2007, Chs 8 and 9) for more on information structuring and discourse strategies in Auslan. Examples 13.5 and 13.6 come from Comrie (1979). Jokes like those recounted throughout this chapter abound on the internet; see also Blake (2007), and Dubinsky and Holcomb (2011). Our approach to cohesion for English follows Halliday and Hasan (1976), and our inspiration for Tables 13.1 and 13.2 came from Czerniewska (1985), Graddol, Cheshire and Swann (1987), and Hudson (1984). An excellent account of how to analyse texts is available in De Beaugrande and Dressler (1996). The wombat names are a reference to the *Muddle-Headed Wombat*; recordings of the radio series are still available for this sweet and silly show, as are the related books by Ruth Park. The version of 'Red Riding Hood' comes from James Finn Garner (1994). For another take on information flow and discourse strategies, see Chapter 9 of Börjars and Burridge (2019). Thanks again to Tania Strahan for sharing her data with us; the transcript included here comes from later in the same conversation as the excerpt in the online exercise relating to Chapter 3. For work on the discourse around transnational human resource management, specifically Filipino 'nurse packaging', see Cleland Silva (2018); the quotation from the recruitment rep appears on page 163. Accounts on discourse markers are plentiful, and a very good place to start is Schiffrin (1987) on *and, oh, well, but, or, so, because, now, then, y'know* and *I mean*; see also Holmes (1988) on *of course* in New Zealand English; Burridge & Florey (2002) on *yeah-no* in Australian English; Makino and Tsutsui (1986), and Puckett (2014) on various Japanese discourse markers; Barske and Golato (2010) on *so* in German; Rezanova and Kogut (2015) on discourse markers in Russian and German science writing. The original paper on narrative structure by Labov and Waletzky (1967) was reprinted in 1997. Cross-cultural differences are reported in Saville-Troike (2003), especially pages 159–60. The research on Aboriginal children's narratives is by Pearce and Stockings (2011) – the quote is from page 130; see also Sharifian (2002), and Sharifian, Rochecouste and Malcolm (2004) for more information on how narratives are structured in Aboriginal English, with text samples mainly from children. Parliamentary records have been used for research on a variety of linguistic features – see, for example, the work by Kruger and Smith (2018). The passage in Mali is taken from Stebbins with the assistance of Tayul (2009). The paper on narrative strategies among Russian-speaking mothers is by Shulova-Piryatinsky and Harkins (2009) (quoted material appears on page 342). The frog story mentioned was originally published as Mayer (1969). A useful orientation to the ways in which children learn to produce narratives is available in Ochs and Schieffelin (2008); see also Andrews, Squire and Tamboukou (2008) on narrative research across a range of settings. The material on English in Japan is from Kubota and McKay (2009) (the quote is from page 594). On the ethnography of communication, see Hymes (1972).

Exercises and discussion points

1 Information structure in Turung

Although Turung is, broadly speaking, an SOV/SV language (recall that we introduced these terms in Chapter 9, in section 9.8 on word order types), pragmatic principles are very often used to organise these constituents as follows: topic, focus, additional argument.

Given this information, how accurate do you think the following English translations are? Are there alternative ways to translate these sentences so that the information structure in Turung is also reflected in the English translations?

[illegible]

[illegible]

'As for the name, I have forgotten.' (Morey, 2010: 221)

[illegible]

[illegible]

‘My mother died when I was small.’

[illegible]*rut*^{3}
[illegible]able
‘When the rice and other dishes are ready, we will eat.’

---

2 On the official record

Following is an example from the public record, this time from Australian Parliament *Hansard* (Question Time, Thursday, 26 November 2009).

Dr Emerson: Now, what I am going to suggest is that the Leader of the Opposition pardons the turkey on the backbench!

Mr Hockey: Mr Speaker …

The Speaker: The member for North Sydney will resume his seat. The minister will either relate his material to the question or sit down.

Dr Emerson: We did make a commitment to the late payment system. We will be issuing a report in the very near future, just after the first anniversary of the commitment coming into force. Of course, we need to report on the first year’s results–

Opposition members interjecting

Dr Emerson: but I tell you this: when it comes out, you will be very disappointed …

(a) Which characteristics of spontaneous speech are missing here?

(b) Why have they been removed?

(c) How many of them do you think were actually present in the original?

(d) What gives you that impression?

3 On interruptions and silence

Interruption is a conversational practice that is valued differently in different cultures and different settings within the same culture. Silence also carries different meanings in different contexts. Notice how you participate in silence and when you interrupt others, then answer the following questions.

Part A: Interruptions

(a) When are you most likely to interrupt someone else?

(b) Who do you interrupt most often? Why?

(c) Who are you unlikely to interrupt?

(d) When do you get upset by people interrupting you?

Part B: Silence

(a) Who are you comfortable spending time with and not talking?

(b) When does silence make you uncomfortable?

(c) Is there any context when you can minimise your response to someone who is speaking to you? Does this really amount to silence?

(d) What does silence mean on the phone?

4 Analysis of a conversation

Examine the following conversation between three housemates talking about media celebrities (from the Griffith Corpus of Spoken Australian English [GCSAusE07]) and answer the questions below.

10. A: [Th]ere's about (0.2) I think three or
11. four times in the (.) the epi↑sode (.) and I was
12. chopping the veggies and Andy was
13. watching it and as it finishes and I am
14. like (0.4) Andy wake up.
15. A: zzzzzzz ((makes snoring sound)) (0.3)
16. M: ha you gotta change the section.
17. A: humm: and I was like (.) I didn't know how
18. to do it and I was chopping
19. so I turned Hamish an Andy on (1.0)
20. driving you home.
21. M: heahherha
22. M: you have the [radio there.
23. L: [They're good aren't they.

24. A: ah they're funny

25. M: hila::rious

26. (1.0)

27. A: ◦(funny kids)◦

28. M: I don't know how they just- and it- if it

29. was just one of them I don't think it

30. would be funny=

31. L: =yeah=

32. A: =no they just (.) bounce off on each

33. other

34. M: yeah

35. L: 'cause Andy actually does most of the

36. talking, (2.0)

(a) Describe the features that identify this language as casual chat (comment on aspects of syntax, vocabulary, meaning and discourse organisation).

(b) Comment on the basic conversational rule sometimes cited in textbooks: 'no gap, no overlap'; in other words, talk must be well orchestrated, where no one is competing for the floor, turn-taking proceeds in an orderly fashion and there are no silences.

5 E-Communication

Collect ten emails and analyse them according to the following questions.

(a) In terms of grammatical structure, in what ways does the language resemble both oral and written language? Give examples.

(b) Do you notice any lexical aspects of interest (for example, non-standard spelling, contractions, abbreviations and word choice)?

(c) How do the writers open and close their messages compared with letter-writing? How have the writers achieved meaning that would normally be conveyed by prosody and paralinguistic features?

Research project

Conversational analysis

Visit the collections in the Australian National Corpus and take extracts of two separate conversations (of around 500 words each).

Outline the various features that distinguish this language as casual conversation (as opposed to writing). Your answer should include aspects such as vocabulary differences, non-fluency features (hesitations, fillers, false starts, self-corrections), grammatical differences and discourse structure (that is, treatment of given or new information and ellipsis). It is a good idea to try to turn the conversations into ordinary written language – it will be a challenge, but you will more readily notice the differences.

Make note of any instances of overlap between the speakers (would you describe these as interruptions or are they more collaborative?). Also note how turns are taken and relinquished, as well as the strategies used to gain the floor and change the topic.

For more activities, please go to www.cambridge.edu.au/academic/linguistics.

14

Language, context and understanding

◆

> When you come right down to it, how many people speak the same language even when they speak the same language?
>
> Russell Hoban (1972), *The Lion of Boaz-Jachin and Jachin-Boaz*

14.0 Introduction

In this chapter, we move beyond the structure of text and interaction to explore the relationship between text, context and meaning. Much of the content is focused on the discipline of **pragmatics**, which is the study of how people make sense out of a message, given the context in which they hear or read it, and the knowledge they have about each other, and about the world and how it works.

Our focus here is on how language and context interact, as well as the strategies used by speakers and hearers, signers, and writers and readers, to negotiate meaning when they communicate. You will see at play the constant competition between economy and informativeness that characterises so much of human interaction. Speakers speak using minimum effort – they do what they have to and they get away with what they can (you saw the phonological fallout of this in Chapter 8). Hearers listen with minimum effort too – they listen with half an ear. (Readers also read selectively – they speculate and read between the lines.) This tug of war between the players, even those of similar linguistic backgrounds, brings into a language many of its indeterminacies and ambiguities.

The writer Cyril Vernon Connolly once suggested that, ‘The perfect use of language is that in which every word carries the meaning that it is intended to, no less and no more.’ If only it were that simple! Even the plainest and most straightforward wording will never guarantee a match between interpretations and intentions. Hearers and readers often attribute to wording all sorts of meanings that were never in the minds of the speakers or writers. As the linguist Fred Householder once observed, ‘nothing can be so clearly and carefully expressed that it cannot be utterly misinterpreted’. Things aren’t as grim as Householder suggests, though – as you will soon see, there are ground rules for engagement that ensure social interaction doesn’t grind to a dysfunctional halt.

14.1 Some preliminary observations

Not everything we say is a simple proposition that can be evaluated as true or false. Language has to cover a huge range of social behaviour – there is so much that goes beyond the literal meaning and that is never subject to precise definition. Imagine the following scenario: a room with an open window, a frail old woman and her young nephew, who has reluctantly come to visit his ageing aunt. Out of the blue, the woman says something like, ‘Do you find it chilly in here?’ She is then surprised because all that happens is that her nephew answers, ‘No.’ The aunt has used an interrogative structure, but her intention was not a yes–no question; it was to make her nephew close the window. She might even describe this situation as her young nephew ‘not doing as he was told’. We will return to this example at a number of points throughout this chapter, developing a framework for understanding what was expected and what went ‘wrong’.

As we saw in Chapter 9, sentence types have conventional associations with particular types of meanings. We generally think that, in English: declaratives make statements; imperatives issue a command; yes–no interrogatives ask yes–no questions; and *wh*-interrogatives ask information questions. However, the relationship between sentence type and meaning type is more like a many-to-many correspondence – each sentence type can express a range of meanings and vice versa. For example, in making a request, English speakers generally minimise any imposition by giving – or at least appearing to give – the option of refusal; they do this by hinting or asking that the hearer do something, rather than ordering them to do it. For this purpose, they have at their disposal a range of different sentence structures that vary in assertiveness and directness.Table 14.1 gives some utterances the aunt might have used to instruct her nephew to close the window.

Table 14.1 Sentence types and meaning

Sentence structure	Example

Declarative	*My goodness, it's chilly in here.*
Declarative	*You could shut the window.*
Interrogative	*Do you find it chilly in here?*
Interrogative	*Are you chilly?*
Interrogative	*Could/would/can you shut the window?*
Interrogative	*Why don't you shut the window?*
Imperative	*Shut the window, dear.*
Hortative (a type of imperative)	*Let's shut the window.*

Languages are full of mismatches between the form of a structure and the meaning. It is what enables them to cope with the complexity of human interaction, but it is also something that can lead to miscommunication. So, it is both a strength and a weakness that languages are not precise notations like logic.

A key focus of pragmatics is to think through how situations like this are handled by participants in the interaction. How is the nephew supposed to know what he is being asked to do? How is he even supposed to identify this question – which appears to be a request for information – as a request for action?

It should be clear, given the 'aunt and the window' example, that effective communication involves much more than parcelling up meaning in a sound signal and broadcasting it through space. For one thing, not only must the listener work at anticipating the expectations and needs of the speaker; the speaker must also take into account the likely assumptions and reactions of the listener. In other words, there is a fairly significant amount of double guessing going on – and effective communication requires a lot of cooperation, too. One way of characterising most of the material in this chapter would be that it theorises and seeks to explain the workings of all this double guessing.

A number of different cognitive strategies have been identified to describe how the listener navigates through the form and superficial meaning of an utterance to arrive at a working hypothesis about the intentions of the speaker. Before we turn to exploring these processes, we want to provide a contrast between how complex interactions are in practice and our common assumptions about how they work.

The conduit metaphor

> Human speech is like a cracked kettle on which we tap crude rhythms for bears to dance to, while we long to make music that will melt the stars.
>
> Gustave Flaubert (1856), *Madame Bovary*

A key part of our discussion here is the idea that speakers and listeners (and, even more so, writers and readers) have to work at communicating meanings to each other. Although people seem to manage to communicate effectively most of the time, this does not necessarily mean that the thoughts, feelings and intentions we set out to share are ever understood by our intended audience in any complete or perfect way. In fact, given the level of cognitive effort that goes into communicating with each other – which is considerable – it is perhaps not surprising that very often things go awry. Sometimes, miscommunication passes without comment, but at other times it changes history.

How can miscommunication have such severe consequences? One reason, Michael Reddy contends, is that the way we talk about language suggests it is a simple case of finding our message and transferring it to the listener. He uses the metaphor of a conduit to illustrate this. (A conduit is basically a tube and language is treated as the tube through which our messages pass.) Most of our common ways of talking about understanding each other conceive of communication as a transfer of goods. Consider the following expressions.

> *Did you get that?* (Has it come out the tube at your end?)
>
> *Can you put that into writing?* (I'd prefer that sort of tube this time).
>
> *That's a loaded argument!* (You're putting a lot of stuff I don't like in the tube!)

None of these expressions reflects the process of deriving meaning from others' utterances as described in this chapter, but Reddy argues that they do create expectations about how communication works and where responsibility lies for miscommunication. According to this model, which we argue is not the best way to describe the process of communication, speakers are generally constructed as the responsible party – they put the message together – while listeners are presented as having a more passive, receptive role.

This model can create problems for people, particularly when the listener deliberately misunderstands and then blames or punishes the speaker for the misconstrued message.

Dangerous metaphors – and misconstrued messages

One of us was once involved in a legal case – *The Medical Board of Australia v Dr D. Roberts*. Here are the details in a nutshell. In a hastily written note, a pediatrician suggested to the mother of twin boys, 'I recommend to your husband that he beat each and any of you our [sic] sons who swear and offend his wife [that is Mother] … to within in [sic] an inch of his life'.

The doctor argued that he was speaking metaphorically to motivate the father to 'take an active role in defending his wife's honour', and he hoped that the 'humour' would help to ease the tensions. The linguist's job in all this was to provide an account of the expression *beat to within an inch of his life.* The sentence illustrates a number of emotive devices commonly used in ordinary language to reinforce the significance of words, so the report focused on two kindred figures of speech: metaphor and hyperbole.

The phrase *beat to within an inch of his life* is one of thousands of English speech formulae structured by a metaphor, and hearers and readers are constantly required to interpret figurative language of this nature. You might compare the fact that, in English, the ordinary way of talking about 'having an argument' is using words like *shoot down, strategy, wipe out* and so on. These expressions draw on a metaphor that speakers of English are barely aware of. Such metaphors, of course, lose their vitality and become hackneyed. For centuries, *within an inch of one's life* has appeared in fixed expressions like *to thrash/beat/flog someone within an inch of their life.* Hundreds of years of use have now reduced this once expressive image to a cliché (or 'lexical zombie', as David Crystal describes it). Here is an example by Lucy Maud Montgomery from 1909 that you will appreciate: 'The grammar class were parsed and analyzed within an inch of their lives.'

Humans are natural-born exaggerators. Phrases like *within an inch of his life* spice up what is being said or written, and make the language more vivid and memorable – and they have been doing so for centuries. True, it is probably not the best wording the doctor could have chosen, but given the circumstances in which the doctor uttered these words, the long relationship between the players (Dr Roberts had been the consultant paediatrician for some years) and the long path of the expression he used, it is hard to imagine how a reasonable person might interpret it literally. Yet, successful use of metaphor and hyperbole, like any other creative language, relies on people cooperating with each other if the communication is going to work – in everyday conversations, insurance policies and messages between doctors and patients.

14.2 Context: Linguistic and situational

When you're lying awake with a dismal headache
And repose is taboo'd by anxiety,
I conceive you may use any language you choose
To indulge in, without impropriety.

(Lord Chancellor) W.S. Gilbert (1882), *Iolanthe,* Act II

One of the important concepts behind much of the way we understand each other is how both the speaker and the listener make reference to the context – the physical and psychological situation in which the interaction is taking place. This is also one of the important skills of good writing: to use the writing to construct a context into which the reader can enter. Communication works most effectively when there is recognition, or the construction, of a shared context. We all go into the act of communication with the hope and expectation that the listener (or reader) will work with us on the task of identifying our intended meanings. This is why a misunderstanding that seems to be deliberate can feel like a betrayal.

I hope you are not going to say I am a sexist misogynist ...

An interesting example of both the gender politics of language and the strategy of blaming the listener for communication breakdown is apparent in a furore that erupted in Australia in October 2014 when the Finance Minister, Mathias Cormann, called the Leader of the Opposition, Bill Shorten, an *economic girlie-man*.

This expression had been used periodically by Arnold Schwarzenegger as a contemptuous accusation of weakness (Schwarzenegger, the action superstar, was also California Governor Schwarzenegger, dubbed 'The Governator'; Mathias Cormann has occasionally been nicknamed 'The Cormanator'!). The term *girlie-man* chimes in with a range of *like a girl* expressions that are intended as put-downs based on physical awkwardness and a lack of strength or ability.

The Australian Finance Minister defended himself from accusations of sexism by claiming that there was no semantic association between the words *girl* and *girlie*: 'I don't think there's anything gender-specific here. Not girls, girlies; it's very different. I hope you are not going to say I am a sexist misogynist.'

We'll leave it to you to go back to Chapter 6 (on semantics), then consider the relationship between *girl* and *girlie* at your leisure. Of interest to us, here, is the suggestion that the listener is being unreasonable in taking offence. Just as it is possible for listeners to deliberately misunderstand the intended meaning of the speaker, it is also possible for the speaker to accuse listeners of doing so.

One way of interpreting the *girlie-man* comment and the ensuing furore is to consider the extent to which the Finance Minister and his audience enjoyed a shared context. The argument about the appropriateness of what he said goes beyond the linguistics of the word *girlie* to whether or not everyone involved agreed on how words relating primarily to women and girls could reasonably be used – a social aspect of the context created in part by the Finance Minister's assumption of public office. In the personal world of the Finance Minister, this was apparently not a point of controversy, though for many people exposed to his public comments it certainly was.

One thing is certain: this round of metadiscourse – talk about the talk – in relation to Corman's statement had the effect of distracting participants from the initial topic. Whatever the audience may have thought of Mathias Corman, Bill Shorten, the physical or economic abilities of girls, or even the hallmarks of masculinity, it is certain that the point Corman claimed he was trying to make was completely overshadowed by the focus on his choice of words and his subsequent attempts to defend them. Whether this represents linguistic misadventure on the part of the Finance Minister or advantage-taking by a group of the more verbally adept is a matter of subjective judgement.

So, how can we come to grips with the idea of context? It is easiest in some ways to think about the **situational context**. In the elderly aunt and nephew scenario, we can explore this with questions such as: Is the window open? Is the elderly aunt lightly dressed? Or, does she have a cold? Yet, situational context is often insufficient to explain all the complexities of how people choose their words or understand what is said to them. The concepts of deixis and anaphora, introduced in Chapter 13, concern the **linguistic context** – when and where the conversation is taking place, and what has already been said. Again, though, context is something richer and even more nuanced than this. It has a social-psychological aspect that we noted above in reference to questions: How much does the aunt believe she should be looked after? Is the nephew feeling kindly towards her? Is she prone to using indirect requests? And what is the relationship between the aunt and her nephew – are they well acquainted, or is this the nephew's first visit?

In the lounge room with one's elderly aunt, sighing over friends and relatives deceased, there is no ambiguity in the sentence *Time flies like an arrow*. But, the same utterance in the hands of someone like Groucho Marx can readily be interpreted along the lines of *Fruit flies like a banana*, suggesting that there is a type of flying insect (the 'time-fly') that enjoys the occasional arrow! Change the situational context and you may readily change the meaning.

In fact, as pragmatists like to note, the type of ambiguity linguists sometimes enjoy exploring in relation to grammar (see Chapter 9 for further examples) is really rather rare in contextually grounded communication. Circumstantial knowledge is often enough to remove the ambiguity and to establish the meaning intended by the speaker or writer. As you might have felt when reading the earlier examples in this book, the experience most of us have is that we zoom in on one sense and are blissfully unaware of any alternative senses. Did you realise, for instance, that *time flies like an arrow* has many more meanings? (Hint: *time* can also be a verb.)

A noteworthy example of genuine ambiguity from the kitchen of a shared house in which one of us used to live occurred when, in discussing possible menu options in the kitchen, someone asked, *Do we have enough* /taem/? The problem was the word /taem/. Talk of food triggered the idea that the reference was to the herb *thyme*, but actually the speaker was worrying about the *time* – how long the recipe would take to cook. The kitchen context happened to make both interpretations reasonable. (Of course, this ambiguity was based on homophony – see Chapter 4 – rather than on syntactic ambiguity.)

As described earlier in relation to the case about the pediatrician, successful use of irony, sarcasm, hyperbole, metaphor and other creative language relies on a rich, complex and, most especially, shared set of understandings about what is happening and what kinds of things would be relevant or appropriate, or even just plain reasonable, to think and to say. One of us can report the embarrassment felt as a small child when her father, as a dinner guest, would exclaim at the end of the meal, 'Well, that was bloody awful!' When people utilise verbal irony in this way, they use words that express a meaning that is the opposite of their literal meaning; there is an incongruity between what might be expected and what actually occurs (usually for humorous or emphatic effect).

Ironic statements (especially in harmonious interaction) typically involve the explicit expression of a positive attitude or evaluation, when the situation actually requires a different – indeed opposite – attitude or evaluation; an example is 'Fabulous day for a picnic', said when looking out the window at a thunderous sky. However, an ironic statement starting with a negative evaluation (such as *Well, that was bloody awful*) is a more dangerous type and is more open to misinterpretation (especially by young onlookers, who only partially share the context!). For speakers to pull off this sort of ironic statement, there must be a certain solidarity and closeness in relationship between the interlocutors, often reinforced by shared context on many levels. While studies have shown that children learn to use and interpret non-literal language earlier than was previously thought (well before primary school, it turns out), there is ongoing research on their developing ability to understand these more complex forms.

Could I have a lawyer?

We've seen that interrogatives are often used for purposes other than posing a question: *Do you have the time?* is a yes–no interrogative, but it doesn't require an answer of *yes* or *no; Why don't you jump in the lake?* is probably not seeking information. In normal conversational interaction, this mismatch between form and meaning doesn't cause any problems because cooperative speakers and hearers take contextual aspects into account (although you will encounter the occasional wag who deliberately answers *yes* or *no* to a request phrased *Do you have the time?*).

But, things can work very differently in a legal context. Diana Eades describes how police officers and courts have disregarded the pragmatic meaning of something like *Could I have a lawyer?* Instead of interpreting it as a request for counsel, they have taken it to be a theoretical query about the availability of counsel. It is an example of what Eades dubs 'the ideology of literalism'. In an ordinary interaction, the implication of *Could I get a lawyer?* would be the wish for access to a lawyer. By appealing to the literal meaning here, these law-enforcement professionals are violating the pragmatic principles that govern everyday exchanges of information between people. Presumably they would not interpret an utterance phrased *Could you pass the salt?* (made at the dinner table) as an inquiry about a person's theoretical ability to pass the salt!

Linguist Anna Wierzbicka attributes the rich abundance of polite directives in English ('whimperatives', as she dubs them) to a culture-related phenomenon that has to do with the emphasis on personal autonomy in Anglo society. Whimperatives like *Could*

I have a lawyer? are more oriented towards the hearer and avoid the impression that the speaker is trying to impose their will on anybody else, as would be conveyed by a bald utterance such as *I want a lawyer* or *Give me a lawyer*. In everyday situations, such utterances are considered abrupt, even rude, and it would be especially absurd to expect a psychologically vulnerable suspect to issue such a directive given the power asymmetry involved in the legal context.

Members of the legal profession seem perfectly able to bypass 'the literal language rule' when it comes to their own courtroom interaction. By convention, they are excessively polite to one other and have ritual face-savers down to a fine art. For example, in Britain or Australia, the formula *In answer to my learned friend's erudite submission* is generally understood to mean 'you are wrong'; in court, criticisms and arguments are always given *with the greatest of respect*. One Melbourne lawyer explained to us that the more uncooperative judges are, the greater the verbal respect accorded to them. In their courtroom doublespeak, lawyers seem quite willing to take account of contextual (pragmatic) aspects of the interaction and 'read between the lines'.

14.3 The force of words

> [I]n all communities, certain words are accepted as potentially creative of acts. You utter a vow or you forge a signature and you may find yourself bound for life to a monastery, a woman or a prison.
>
> Bronislaw Malinowski (1935), *The Language of Magic and Gardening*

One way of picking apart the speaker's meaning from their words is to consider the force behind their utterance. The challenge here is to be able to identify times when the words are said with the intention of making specific, and perhaps lasting, changes in the world as opposed to those times when things are said with less definite intention. Utterances that 'mean business' are known as **speech acts**. They include requests, promises, apologies, predictions, warnings, invitations, resignations, assertions and so on. So, think of 'speech act' as a special term to describe those sentences that have a performative function – in other words, they are defined in terms of the intentions of the speaker and the effect they have on the listener. Although speech acts have been described as staple items of communication, it wasn't until British philosopher J.L. Austin's *How to Do Things with Words* that 'speech act theory' really took off. In this work, Austin made a three-way distinction between different types of actions associated with speech acts.

1 The act of performing the utterance – called the **locutionary act** (*locution* refers to the form of an expression). The elderly aunt says: *Do you find it chilly in here?*

2 The social act resulting from the utterance – called the **illocutionary** act (*illocution* refers to the action intended with the utterance). The elderly aunt (thinks she has) made a request for her nephew to shut the window.

3 The effect of the utterance – the **perlocutionary act** (*perlocution* refers to the results of the utterance – the way the listener is expected to respond to them). The dutiful nephew is expected to shut the window (and, furthermore, fetch a quilt) for the aunt.

The focus of much writing on speech acts is on illocutionary force – what the speaker intends to accomplish by making the utterance.

The relationship between words and actions is particularly clear with one special set of speech acts called **performatives** – promising, pronouncing, baptising, conferring and betting are types of speech acts whose illocutionary force is consistent and clear. Provided the speaker is empowered to do the action and any other conventional requirements are in place, the act is accomplished by uttering the words; for example, *I assert that Cosmo stole the sock*. So, performatives are types of sentences that are equivalent to actions. You can test whether a sentence is a performative with something called the 'hereby test'. Try inserting the word *hereby* in the sentence (*I hereby assert that Cosmo stole the sock*); sentences that aren't performatives will sound funny (?*I hereby know that Cosmo stole the sock*).

Indirect speech acts are those where the speaker's meaning, intention or desired outcome is not directly encoded in the utterance. Making the distinction between what is said and what is intended allows us to come to grips with all sorts of indirect ways of communicating. *Do you find it chilly in here*? uttered as a request for action is a classic example of an indirect speech act.

The main reason perlocutionary acts are proposed is to acknowledge that what we say (directly or not) may not have the consequences we intend. To get the results the speaker wants, the hearer has to recognise what is going on and comply. This doesn't always happen. A simple example is to consider the expression *How are you?* This is uttered, in most cases, as a greeting rather than a request for information. Occasionally, the conventional interpretation is not the intended one (the speaker really does want to know, in detail, how the listener is feeling) or the response from the listener is incorrect (there is a mismatch between how they interpret the question and what the speaker intended). For example, they may provide extensive detail on their physical wellbeing, when all that was really being sought was a grunt and a nod of the head.

14.4 The cooperative principle

Given the variability and indeterminacy of meaning, and the mismatch between linguistic forms and their meanings, if communication is ever going to work, interlocutors have to cooperate with each other. As with any other activity, a linguistic exchange can only take place successfully if all the players involved mutually recognise that a set of ground rules governs their use of language and also their interpretations of the use of language by others (consider, for example, what is assumed and required in relation to insurance policies, doctors' communications with patients and everyday conversations). A possible set of ground rules was outlined by Paul Grice in 1975; these are known collectively as the **cooperative principle**. They can be summarised in the following terms: 'Make your conversational contribution such as is required, at the stage at which it occurs, by the accepted purpose or direction of the talk exchange in which you are engaged.' Grice argued that listeners used the cooperative principle to make sense of what speakers intend by their utterances, and he unpacked it into a series of maxims (more specific rules or conventions). Since the time Grice first proposed the cooperative principle, other researchers have considered how it applies (or fails to apply) in different cultures. We will explore this topic in more detail later in the chapter.

Maxim of quantity

The maxim of quantity is all about determining how much to say (or expect the speaker to say). It is presented in two parts.

1 Make your contribution as informative as is required (for the current purposes of the exchange).

2 Do not make your contribution more informative than is required.

In other words, don't say too much or too little – just enough to make your message clear. For example, you normally wouldn't say, 'My poodle, which is a dog, is having pups.' The clause 'which is a dog' is totally unnecessary – if something is described as a poodle, then the hearer can safely assume that it is a dog.

When it’s best not to tell

The maxim of quantity presupposes some sort of access to knowledge as a right – that the normal response to any kind of request for information (direct or indirect) should be met with compliant provision of the information. Academic research in the Western world seems to run on this assumption – often expressed along the lines that knowledge (education and research) is good and valuable, and should be shared with everyone. In contrast, knowledge in many Indigenous communities is held under a model of custodianship. It is not freely shared; rather, it is protected, managed and passed along only to appropriate people with the capacity to understand and care for it in turn.

In linguistics and other related disciplines, such as anthropology, fieldwork relies on the researcher being able to obtain information about the community they are studying. However, not all communities of interest to researchers hold this ‘access to knowledge as a right’ perspective, and difficulties can arise for both researchers and the communities themselves when these two different values systems meet. In some places, people may not feel that the maxim of quantity – especially ‘Make your contribution as informative as is required by the researcher’ – is a good idea.

Tonya’s research in the Mali community is a case in point. The Malis (and more generally the Baining ethnic group to which they belong) have a somewhat notorious reputation as obstructive subjects based on their refusal to divulge the details of their sacred dances to outsiders. One prominent anthropologist went so far as to say that they broke his heart. On one level not complying with requests for knowledge may be expressed by a lack of cooperation in a range of practical ways; for example, by people refusing the researcher access to items or activities of interest. On another level, though, it can become a quite specific communication problem – one that can be understood with reference to the maxim of quantity.

Responses to direct questions that fail to provide expected levels of information may or may not reflect the level of knowledge of the speaker, and this depends partly on how the cooperative principle is operating (if it is indeed operating at all), as well as how participants from different cultures relate to one another and to the maxim of quantity. Tonya certainly felt, as a woman for whom much knowledge was taboo, that this maxim in operation for her had particular meanings attached to the phrase *as required*. She was happy to work in the community focusing on non-secret topics – in practice, *as required* was replaced by *as appropriate according to Mali assumptions about the safe sharing of knowledge with someone like Tonya*.

Maxim of quality

The maxim of quality is focused on describing how the listener is able to rely on what the speaker is saying. It has a 'supermaxim': *Try to make your contribution one that is true*. Again the maxim is divided into two parts.

1 Do not say what you believe to be false.

2 Do not say that for which you lack adequate evidence.

In other words, you should be genuine and sincere – state as facts only what you believe to be facts; make offers and promises only if you intend carrying them out; pronounce judgements only if you are in a position to judge and so on. For example, if we told you that the chapter 'Logic and Conversation' in Grice's 1975 text would make great bedtime reading for five-year-olds, we would clearly be violating this maxim!

Telling alternative stories

Again, there are cultural nuances associated with this maxim. One of the problems for researchers and consultants working across the secret–sacred boundary is that very little that is true can be said openly. This reflects the dilemma of secrecy – as Hugh B. Urban observes, 'if one "knows", one cannot speak; and if one speaks, one must not really "know"'. Secret teachings require initiation, but once obtained they cannot then be communicated to non-initiated outsiders. In other words, having been initiated, you cannot say that for which you have adequate evidence.

In the Mali community, there are many other situations in which people routinely do not provide information that they believe to be true. Here is what Tonya and her co-researcher

Mark Planigale wrote on this:

> Gricean maxims do not apply universally … In the Mali context, there is an entrenched commitment to privacy that can seem like secrecy to an outsider. In Marunga people often leave the village discreetly and routinely give misinformation about where they are going, even when they are simply off to their garden for an afternoon of work. While to many people this behaviour might be interpreted as frustrating or even offensive, in Marunga it is an enactment of good community membership.

Maxim of relation

The maxim of relation is intended to provide the means for dealing with a range of phenomena associated with the ways in which people wind their way through conversational topics. Grice acknowledged that it ‘conceals a number of problems’ and indicated that it was intended as something of a placeholder. The maxim is:

Be relevant.

In other words, utterances should have some bearing on the context in which they are uttered. Relevance is understood to be of such importance that it forms the basis for another theory of communication. **Relevance theory**, developed by Dan Sperber and Deidre Wilson, elevates relevance to the central principle: the communicative principle of relevance. This principle works by assuming that the listener’s main strategy for arriving at the meaning intended by the speaker is to assume that what has been said is intended to be relevant to them. Thus, the question *Do you find it chilly in here?* addressed to a young person with excellent circulation and plenty of clothes on is more likely to be relevant as an indirect request than an inquiry after their welfare.

Maxim of manner

The maxim of manner has the most enjoyable supermaxim: *Be perspicuous*. (Perspicuous means clear – it is perversely pleasing to us that the supermaxim seems to flout the maxim itself for many readers – including us the first time around … and Grice does it again under point 3 below.)

1 Avoid obscurity of expression.

2 Avoid ambiguity.

3 Be brief (avoid unnecessary prolixity).

4 Be orderly.

In other words, be clear, be organised and don't be long-winded – don't mislead or confuse your hearers. For example, you wouldn't normally refer to a toothbrush as a *home plaque removal implement* or potholes in a road as *pavement deficiencies*. This would be engaging in excessive wordiness – or prolixity (which our spell-checker keeps changing to proximity!). In other words, if you're not presenting your meaning in a clear and concise manner, you are causing your hearers undue mental effort.

Does your dog bite?

Comedy routinely derives its humour from not adhering to the cooperative principle. Below is the classic 'Does your dog bite?' joke followed by some other classic (and rather corny) jokes. All show the characteristic lack of cooperation that we expect in this sort of humour. For each, see whether you can figure out which maxim has been contravened.

(1)

Man: Does your dog bite?

Woman: No.

(Man reaches down to pet the dog. The dog bites man's hand).

Man: Ouch! Hey! You said your dog doesn't bite.

Woman: He doesn't. But that's not my dog.

(2)

Barrister: Doctor, how many caesarians have you performed on pregnant women?

Doctor: All my caesarians have been performed on pregnant women.

(3) As he paid his hotel bill, the guest turned and called to the bellboy, 'Quick, run up to room 225 and see whether I left my briefcase. Hurry, please, because I've got just five minutes to catch my train'. Three minutes later, the bellboy was back, all out of breath. 'Yes, sir,' he gasped, 'it's there all right.'

(4)

Question: Excuse me, do you know what time it is?

Answer: Yes!

(5)

Question: How many surrealists does it take to screw in a light bulb?

Answer: Fish!

(6)

Recruit:A drowsy numbness pains my sense.

Drillsergeant:Who said that?

Recruit: Keats, wasn't it?

Do not think of Grice's maxims of cooperative behaviour as laws that must be obeyed. They are reference points for making sense of language in action. Linguist Keith Allan suggests that a useful analogy is the compass, the points of which provide a conventional framework for giving directions and specifying locations over the surface of the Earth. If there was no cooperative principle, systematic communication would be impossible. Speakers would have no ground rules for getting their message across to the audience and, conversely, there would be no basis for deciding whether utterances made sense or what value should be put on them.

14.5 Extracting meaning from talk

> When a man talks rot, that's epigram: when he talks sense, then I agree with him.
>
> George Bernard Shaw (1908), *Getting Married*, Act I

So how do you extract the speaker's intended meaning from what is said? Consider the following interaction (first introduced in Chapter 13):

(14.1) [Doorbell rings.]
Fritz:That's the door!
Siggy:I'm in the shower!

Fritz: Okay.

At face value, these sentences do not seem to amount to a meaningful exchange, yet both Fritz and Siggy understand exactly what is going on. Siggy implies (makes it understood without expressing it directly) that he can't answer the door. Fritz infers (makes a reasonable guess) from this that he needs to answer the door. Inference and implicature are two sides of the same coin; speakers typically imply more than they assert and hearers typically infer more than is asserted. This sort of thing is going on all the time in verbal exchanges, especially in casual conversations where so much is assumed, taken for granted and, in the case of good mates, there much common ground.

Contravening maxims

Most interaction operates with the notion of harmony in mind. All participants normally assume that others are providing the right amount of information, telling the truth, and being relevant and as clear as possible. Even if a speaker appears to violate a maxim, we do our best to interpret the violation as meaningful. So, if someone did say to you, 'My poodle, which is a dog, is having pups', you would probably do your best to figure out why the person had included such a piece of apparently superfluous information. Here are some other scenarios that illustrate this point.

- *Scenario 1:* You have just asked a friend how the family party went earlier in the week. Your friend replies: 'Oh, you know, it was a family party.' At first blush, this is not a particularly informative comment! But, you expect that your interlocutor is being cooperative and communicating something. The response suggests that it was a typical family party – that seemingly unhelpful statement speaks volumes and there's not much more to be said.
- *Scenario 2:* In the middle of a conversation about the football grand final, your conversational partner suddenly announces she is going overseas at the end of August. At first, this appears an irrelevant remark to make. Yet, we assume that this person has a reason for making the comment. If she is being cooperative, then it must somehow be appropriate to the discussion. So, you will do your utmost to make sense of it. You might infer, for example, that she can't attend the

grand final match in September.

Even failing to follow the discussed maxims in cultures where they normally apply can often be understood as a type of cooperation in some circumstances. In the educational setting, for instance, instructors routinely ask students to give information that is already well known to them (in other words, in breach of the maxim of quantity). The same would be true of courtroom exchanges, where witnesses are called on to provide answers that may already be known. In addition, non-literal uses of language (such as irony, sarcasm, hyperbole and metaphor) flout at least one of these maxims but can generally be understood with reference to the others and to the overarching principle. Consider the following examples of non-literal language and try to determine which of the maxims is being flouted here.

- *Irony:* an expression that normally signifies the opposite to the literal meaning; for example, an inappropriately formal register might be employed to express irony, as in *A good time was had by all.* (Sometimes a sentence-final *not* makes it clear that the speaker was being ironic *A good time was had by all … NOT!*).
- *Sarcasm:* an ironic remark that appears to be in praise of someone or something but is in fact taunting or hurtful; an example is *wonderful work* when said of something that is clearly woeful, or something bad happens (the top has come off the blender, spraying beetroot around the kitchen) and it provokes the response *Terrific, that's just what I needed!*
- *Hyperbole:* exaggerated language often for the purpose of emphasis (there are millions of reasons for this!); for instance, it is often used to magnify offensive intent (*You rotten bloody mongrel bastard!*) or to diminish and downgrade a person (*You pea-brain!; You slimy little toad!*).
- *Metaphor:* the use of figurative expressions (as in the examples of *pea brain* and *toad* just given) to refer to something that they do not literally denote in order to suggest a similarity; for instance, someone might be described as having *raven hair, ruby lips, emerald eyes, teeth of pearl* or, perhaps, even a *red cabbage nose.*

In our everyday conversations, Grice's maxims are routinely contravened. In the above examples, speakers choose to ignore (or flout) the maxim of quality. But, they do not want to deceive; they rely instead on the listener's ability to add meaning beyond the literal meaning – the listener is expected to look for extra significance. In the last example, the metaphor ‘red cabbage nose’ is a lie unless that person is some kind of Frankensteinish monster!

In addition to flouting maxims for various communicative purposes, Grice acknowledges there are many ways that speakers may fail to fulfil a maxim. The following examples will give you a better idea of how non-observance can work.

- Violating a maxim involves deception of some sort. Hearers are not encouraged to look for implicatures, but rather to take the utterances at face value. This could include something like a white lie; for example, in response to the offer *Do you want to come out for a meal?* a speaker might violate the maxim of quality and lie about a previous engagement (the lie saves the situation by indicating that the relationship is still where it was).
- Speakers might inadvertently disregard language errors resulting from defective performance; this might involve a child or adult learner of the language, or perhaps even an intoxicated talker.
- Speakers might also opt out of these maxims if they need to withhold information for ethical or privacy reasons – as, for example, in cases of doctor–patient or lawyer–client confidentiality, or the patterns of (not) sharing information in the example of Mali mentioned earlier. This would also involve situations where the maxims are suspended; for instance, a participant might withhold information for cultural reasons (it might be culturally inappropriate to provide the information, so the maxim is suspended, as again in the Mali situation).

We now will review three different ways in which indirect meanings are communicated by speakers and understood by hearers: presupposition, inference and implication.

Presuppositions

One important strategy that listeners can use to cut down on possible interpretations of utterances is to consider what they could treat as presupposed by the utterance or text. Presuppositions allow the listener to narrow down likely meanings of an utterance. For example, *Where is Maisy living these days?* presupposes that there is someone uniquely identified as Maisy known to the addressee as well as the speaker, and this person has moved. It may not be true that Maisy has moved, but the speaker's belief that Maisy has moved is built into the question. Presuppositions are applied by the listener to the content of the utterance or text in order to figure out what is meant. Thus, *Where is Maisy living these days*? provides the listener with the presuppositions that the speaker does not know where Maisy is living and that the speaker thinks that she has moved, and that the listener might know where she has moved to. These presuppositions assist the listener in formulating responses, which may include responses like *Oh, she's still at her parents' place for the time being*. So, presuppositions are the background assumptions that both speakers and hearers must share for utterances like this one to be considered appropriate in the given context. Similarly, anything said by Fritz in response to the doorbell ringing is likely to involve the idea that someone (perhaps Siggy) should respond to see who is there.

Inferencing

Understanding each other involves a process of thinking through what has been said and considering other relevant information that can assist the listener in reaching a reasonable interpretation of the meaning under the circumstances. This process is known as **inferencing**. Inference-making by the listener is one of the significant omissions of the conduit metaphor. When one's aunt comments on the temperature in the room, in formulating a response one may consider her age, her level of mobility, her general health, her understanding of how and how much she should be looked after, whose home it is, whether one is feeling kindly towards her, whether she is prone to using indirect requests and so forth. All this inferencing in relation to the utterance *Do you find it chilly in here?* allows the listener to arrive at an appropriate response.

There are two particular types of thinking that have been identified as relevant to the process of inference: the identification of **presuppositions** (things that speakers assume (or act as if they assume) in order for their utterances to be meaningful, as described above); and the identification of **implicatures** (things implied by what the speaker says). We hasten to add that these concepts are difficult to tease apart and controversial – not all would agree with our account here.

Implicature

The cooperative principle describes a set of assumptions followed by the speaker as they work out how to communicate their message and applied by the listener in trying to understand what the speaker is saying. We have already shown how the specific detail or the workings of the cooperative principle may vary depending on the cultural context, and we return to this point below. Here we consider how speakers and hearers arrive at reasonable understandings of what is being said when the cooperative principle superficially appears to have been ignored.

Grice argues that when this occurs, the hearer will assume that the speaker is apparently flouting a conversational maxim because the speaker is communicating in an indirect way. In other words, the cooperative principle is still expected to apply, but the hearer has to look for less obvious meanings in order to understand how the speaker's words are consistent with the cooperative principle.

For example, when Fritz observes that there is someone at the door, Siggy might well wonder what that has to do with him. The observation would be flouting the maxim of relation (that is, be relevant) unless the presence of someone at the door is associated with an expected response of some sort (which of course it is). Fritz is using his utterance to open negotiations about who is going to answer the door. The conversational implicature (the thing he has not said directly) is that he is hoping Siggy will do it so he doesn't have to.

When Siggy says 'I'm in the shower', it is Fritz's job to work out how this could be a relevant and appropriate response. This response also appears to flout the maxim of relation. Assuming the cooperative principle applies, Fritz is likely to realise that Siggy being wet and naked in the shower is not compatible with Siggy answering the door, meaning that Siggy's response is, in fact, totally relevant. Fritz is going to have to do it himself.

Similarly, if someone was asked *Where is Maisy living these days?* and replied *Somewhere out west*, they would seem to be flouting the maxim of quantity. Any one of a number of more explicit answers would logically be acceptable and even desirable, given the cooperative principle. Maisy could be living in a distant suburb in the same city, in another state or on the other side of the country. A vague answer to a specific question often implies that the speaker cannot supply more detailed information. Not providing more specific information when it is available might seem unhelpful or even rude unless they had no further information to give. Applying the cooperative principle allows the hearer to understand that this must be the case.

As we noted above, there are many cultures in which knowledge is shared in sometimes quite closely managed ways, and it is not particularly reasonable – especially in interactions between young people and much older people – to expect that complete answers will routinely be given. Vague answers from an older person could just as well be intended to communicate conversational implicatures such as, *I'll tell you once you have a bit more life experience under your belt* or *There are a few other specific things I need to get around to telling you first.*

14.6 (Im)politeness

Politeness is another reason why the cooperative principle is flouted regularly. In many English-speaking societies, for example, there is a strong correlation between the length and indirectness of an utterance, tight and its supposed level of politeness. Consider the following ways of accomplishing the task of getting someone to shut the window:

(14.2) Shut the window.
Will you shut the window?
Can you shut the window?
Would you mind shutting the window?
Could you possibly shut the window?
It's a bit chilly in here.
Do you find it chilly in here?

Indirectness is interesting because some cultures view it as an underhanded way of getting what you want (and hence would find the *bit chilly in here* strategy manipulative), while others see it as a means of avoiding an expression of power over others (*We both know I can tell you what to do but let's pretend you're a free agent …*).

Changing pronouns to reduce directness

As we have seen, there are a number of ways to change the directness of a command. In English, we tend to go for converting an imperative (the more conventional form to request action) into an interrogative, presenting our command as a question. In the Australian language Gurindji (spoken in the North Territory), rather than changing the clause type, the form of the pronoun is changed to increase the distance between the speaker and the implied command. (You might compare how, in rather stilted English, hearers can be addressed in the third person. You'll be familiar with *If Madam so desires, she could have our tailor alter the waistband just a little*.)

See examples 14.3 and 14.4. Both are intended as commands and are directed to the listener, but, in the more polite version, the third-person pronoun 'those two' is used instead of the more direct second-person pronoun 'you two.'

(14.3)

Nangari-wu	Marnta-wula-rla
skin.name-to	talk-those.two-to.her

You two talk to Nangari! (Literally: Those two talk to Nangari.)

(14.4)

Nangari-wu	Marnta-nyula-rla
skin.name-to	talk-you.two-to.her

You two talk to Nangari!

Example 14.3 is the less-direct form and is the usual way to ask someone to perform an action. Example 14.4 is considered impolite speech, and examples of second-person pronouns in commands are rare in Gurindji interaction.

Geoffrey Leech developed the following six maxims of politeness; these provide us with a useful catalogue of the types of behaviours that are likely to cause trouble in terms of politeness.

1 *The tact maxim:* don't say things that will upset other people.

2 *The generosity maxim:* say things that imply or cause others to get the better deal.

3 *The approbation maxim:* avoid disagreement and try to say things that make others feel good.

4 *The modesty maxim:* praise other people and put yourself down.

5 *The agreement maxim:* minimise expressions of disagreement.

6 *The sympathy maxim:* express sympathy to others; do not express dislike.

As with our example of making a request, these maxims do not work in the same way across all cultures, a topic to which we will return shortly. First, though, we want to introduce you to the concept of **face**.

Face

The concept of face, particularly as it is construed in the expression *to lose face*, is central to much work on politeness. Face is understood as the need people have to be well regarded – it relates to feelings of self-esteem and to the social status associated with one's reputation or good name. (This use of the expression *face* came into English from Chinese, and is current in everyday parlance, especially in figures of speech such as 'to lose face' and 'to save face'). Work by Erving Goffman established the concept of face in the social sciences, and Brown and Levinson expanded the term by considering different aspects of face.

Brown and Levinson identified two different aspects of face. The first is our need to be well regarded (called positive face) and the second is our need to have the freedom to act as we wish (negative face). Goffman, and then Brown and Levinson, also considered how face could be maintained: Goffman by describing **facework** as efforts associated with maintaining one's own and others' face; Brown and Levinson by considering how people manage face-threatening acts. A **face-threatening act** is one that either conveys dislike of the listener (threatening their positive face) or imposes some demand on them (threatening their negative face). The following table sets out a number of examples and shows how things the speaker says or does can be a threat to their own face as well as to the face of the hearer. As we have already noted, in practice, these observations may vary in different settings and cultures.

Table 14.2 Threatening the face of hearers and listeners

Threat to hearer	**Threat to speaker**
Commanding, requesting (negative face – the threat involves the creation of an obligation on the part of the hearer)	Thanking, accepting an offer, accepting an apology
Interrupting, boasting (positive face – the threat involves experiencing lack of regard or lack of care)	Apologizing, confessing, losing control

In order to decide how to manage face-threatening acts, Brown and Levinson propose that the speaker needs to balance several different issues against each other: social distance, relative power and a ranking of impositions (because, after all, taking out the rubbish is a lesser task – a smaller imposition – than repainting the whole house).

Social distance

Social distance is all about familiarity. We typically behave differently with people who we know well and see regularly than people with whom we are less familiar. Some researchers have argued that it is not so much the contrast between intimate associates and very distant others that is relevant here; instead, they suggest we need to do more facework with people in the middle of the scale of social distance – we have to work harder to negotiate good relationships with people who are not complete strangers, but who are not intimates either. Thus, asking your partner for a lift home from work is a different proposition from asking a workmate (although the duration of acquaintance with the workmate is also likely to have an impact on how the request is made).

Relative power

Relative power refers to differences between people in terms of how they are able to assert themselves in relation to each other. Asymmetrical relationships involve people who differ in their ability to make the other person do something or agree with what they have to say. From an Antipodean perspective, this is a bit of a challenging topic because Australians and New Zealanders share the trait of being resistant to the idea that there are power differences between people. Nevertheless, such differences do exist in practice. Asking your boss for a lift home from work is different from asking your partner or a colleague.

Ranking impositions

Some things that we ask of each other are more significant than others. Our contrast between taking out the rubbish and painting the house is a simple case in point. Brown and Levinson note that there is a level of cultural nuance in the way similar impositions may be viewed in different cultures. In other words, there are lots of potential subtleties here – for example, asking someone to take a message to their son-in-law has vastly different implications if this relationship is taboo. In Mali society, the son-in-law is rarely addressed directly and does not eat or drink in the presence of his parents-in-law. Furthermore, there is a name taboo in place that means that even referring to the son-in-law causes distress to the father-in-law or mother-in-law. Even framing such a request amounts to a significant face-threatening act in this setting.

Face on Facebook

In this book, we have spent a lot of words talking about language as a mirror of social attitudes, personal associations and prejudices of individual speakers. Online communities offer another, rather disturbing, perspective on politeness phenomena and conventions for language behaviour.

Monash University student Amanda Young examined the topic of racism for her honours thesis; in particular, racism towards immigrants and asylum seekers in Australia. Her study followed various Facebook groups over a twelve-week period and analysed the language according to politeness markers, euphemisms and dysphemisms, and the nature of inflammatory language. What she found was that away from face-to-face contact and within the safe confines of their internet personas, people are much more likely to dispense with politeness paraphernalia and express racist views – and in the most inflammatory language. Her thesis captured the consequences of a virtual community where people can ‘say’ whatever they want whenever they want. As she concludes, the shield of anonymity that being online offers people is a way of bringing to the surface all those gritty things that politeness, and even just plain civil society, would prefer to stay hidden. There is also the chance that some of these profiles were created specifically so a person could voice their own opinion without anybody from their real life knowing about it. Facebook offers an interesting experience in that it is both private and public.

14.7 Cross-cultural pragmatics

One of the challenges to theories of politeness is the way the parameters described so far vary from one culture to another. Matsumoto discusses the notion of face in Japanese, arguing that the concept is qualitatively different from the way it has been constructed by Brown and Levinson (to reflect preoccupations with individualism in the West). Matsumoto characterises Western-style face as being about maintaining individual territory (that is, status and freedoms). She notes that:

> What is of paramount concern to a Japanese is not his/her own territory,

> but the position in relation to the others in the group and his/her acceptance by those others. Loss of face is associated with the perception by others that one has not comprehended and acknowledged the structure and hierarchy of the group.

Matsumoto goes on to contrast conventionalised expressions of politeness associated with gift-giving in different cultures. Whereas the sentences in examples 14.5 and 14.6 are presented as textbook examples of polite offerings in Japanese, Matsumoto notes that example 14.7 is heard by Japanese speakers as rather rude because it implies a lack of taste on the part of the receiver.

(14.5) *Anoo, tumaranai mono desu ga …*
(Lit.) ‘Um, this is a trifling thing, but … ’

(14.6) *Okuti ni awanai kamosiremasen ga, ohitotu doozo*.
(Lit.) ‘This may not suit your palate, but please accept just one.’

(14.7) *It’s not much, it’s just a little thing I picked up for a song in a bargain basement sale in Macy’s last week, I thought maybe you could use it.*

According to Matsumoto, the fact that evidence is provided for the gift being cheap is what creates the feeling of rudeness in example 14.7, whereas the assertion that the gift, alone, is insignificant, as in 14.5, does not have this effect. There is also a noteworthy difference in the goals of the two utterances. Example 14.7 is uttered as a way of reducing obligation on the part of the recipient (*Don’t thank me, it’s not a big thing and I’ve provided you with evidence that this is the case*), while the motivation for example 14.6 is to establish that the recipient has excellent taste (*I really don’t know if you’ll like this but …*). Matsomoto comments: ‘In terms of Japanese society, however, freedom from imposition is far less potent as a dynamic of politeness than is preservation of the social ranking.’

With globalisation comes an increasing need for intercultural communication, and as researchers such as Ogiermann and Wierzbicka point out, we need to find out more about what is and is not polite across different cultures.

Saying what you mean in Russian

Indirectness was established by Brown and Levinson as an important resource for managing difficult communications. Basically, they proposed a direct relationship between increasing levels of indirectness and politeness – the more threatening the overall intention, the more facework, in the form of indirectness, has to be done. For example, refusals and denials of any sort are potentially hazardous to face and, as we saw earlier, speakers making a request can minimise an imposition by giving – or at least appearing to give – the option of refusal by hinting or asking that hearers do something, rather than ordering that they do it (*Do you think … ? Would you … ?*).

It has become increasingly clear, however, that the value of indirectness as a politeness strategy is closely bound up with Anglo-Saxon culture (among others, including Gurindji, as noted earlier) rather than being universal. For some cultures, directness is associated with openness and honesty. Far from being polite, indirectness can be seen as devious and manipulative – even a waste of time for hearers who have to read between the lines!

Hence, requesting behaviour can be very different around the world. Speakers of Slavic cultural backgrounds don't regard a request as an imposition on a hearer's personal freedom and the threat of a refusal doesn't hold the same potential loss of face. The use of imperatives in making requests is seen as quite appropriate in languages like Russian; however, there are strategies in the construction for showing consideration towards the hearer – for example, different verb forms can capture nuances of politeness and friendliness. In Russian, the use of the imperfective aspect in a command is polite and inviting, while the perfective is more strongly commanding and the infinitive is compelling (see Chapter 5 for a discussion of the grammatical category of aspect).

Silence across cultures

Interactions are hard work. Talk must be sequenced, and it must be opened and closed effectively. People can go about this in very different ways. There is also the problem of when to talk, how much to talk and when to remain silent. The most successful communication will usually be when speakers have the same habits and attitudes, especially when it comes to the expectation of silence. But, no group will be homogenous – even individuals within the one speech community can have vastly different interactional styles. Some speakers are noisy or show enthusiastic involvement by way of questions and overlapping comments (what linguist Deborah Tannen has dubbed a 'high-involvement style'). Other speakers seem to almost wait for the floor to be empty before they take their turn; they would never dream of imposing on another speaker (Tannen's 'high-considerateness style'). Things don't always go smoothly when speakers with very different interactional styles get together. From the perspective of high-considerateness, an interruption suggests you don't care, you're probably not listening, perhaps you're not even interested. But, this may not be the case. On the other hand, enthusiastic overlappers may feel hurt by a lack of participation because it suggests to them a lack of interest or support. Is an interruption an act of verbal cuddling or an act of verbal aggression?

As you might imagine, these differences in interactional styles can have serious consequences for how people relate to one another, especially in cross-cultural communication. Consider how differently silence is interpreted across speech communities. English speakers attach a high value to talk and silences are usually viewed negatively. Even when they indicate uncertainty or reflection, they can be awkward and uncomfortable, but often they are viewed as 'loud' indicators of hostility or distress. In other communities, it can be very different.

Consider the Pennsylvania German-speaking Anabaptist communities. Now, we don't want to give the impression that these speakers don't love to chat, but silences are commonplace and comfortable. Communication doesn't have to be verbal, and visiting close friends and family can involve the exchange of very few words – people can happily sit in silence for long periods of time, as Kate (who would otherwise keep talking at all costs!) has experienced. In this community, speech doesn't have the same sort of interpersonal function as it does in, say, mainstream North American culture. There is silent prayer, tolerance of long between-turn silences, and also patterns of 'pleases', 'thank yous', greetings and leave-takings where silence is commonplace. There is often quiet, for instance, at the dinner table when food is being passed around, and while

Pennsylvania German does have greeting formulae, it is also usual for speakers to say nothing. In such an integrated community, where people are deeply involved with one another and where there is no social distance, there is not the same requirement for language to express information about values and social relationships as there is in the more mainstream culture. Quite simply, such things can be taken for granted in this community. It is difficult for someone not native to the culture or language to recognise what is really going on when this sort of interaction takes place.

In Chapter 3, we mentioned that one of the difficulties of undertaking linguistic work is the observer's paradox (observing the way people use language when they are not being observed). Linguists aim to collect natural language, but their presence destroys this possibility. Observing moments of silence is particularly problematic, and linguists from outside the community may never be in a position to observe it, since their presence will always disrupt the natural social setting that gives rise to these silences.

A key skill that Tonya developed in her fieldwork and now uses in a range of other settings has been a confidence about letting silence come to its own fruition by waiting for others to speak. In the early days, she realised that she needed to give others more room to take their turns, and often resorted to sitting on her hands to manage the urge to fill silences with encouraging talk and questions. What people really needed was time to gather their thoughts, and space that allowed them to opt in to an active role in the interaction. Silence was far more effective than chatter!

Silence is talk

People in Australian Aboriginal communities are often described as being comfortable with lengthy silences. Given that many of these communities, particularly in remote areas, consist of people who, like the speakers of Pennsylvania German, are deeply involved in one another's lives, with shared knowledge and clearly defined relationships, silence may well be comfortable for similar reasons. Ilana Mushin and Rod Gardner set out to explore the ways in which silence operates in one remote community, putting some data around the more informal observations that others had made. They found that, compared with data from languages such as English, silences extending for longer than one second were more common and were less likely to be associated with interactional problems such as word searches, confusion or disagreement. The comfortableness of these silences is partly shown by the ease at which talk resumes – it appears as if the silence has no impact on the way conversations proceed. There are also differences in comparison with English in the way silence is used in allocating turns between speakers. Overall, these differences indicate less pressure to take the floor from a previous speaker as soon as it becomes available. These different norms about participation, particularly in relation to turn-taking, deserve attention, especially, in institutional settings where particular patterns of interaction are expected, and Aboriginal people may be disadvantaged if cultural differences in relation to their participation are not recognised.

When insults aren't insults: The 'banter principle'

First youth:*Hullo, congenital idiot!*

Second youth:*Hullo, you priceless old ass!*

Damsel:*I'd no idea you two knew each other so well!*

You will be familiar with the kind of scenario illustrated above in an old *Punch* cartoon from the early 1900s. Here you see the exchange of abusive terms with intended meanings that are the complete opposite of their primary meanings. It's a kind of mock impoliteness that is common between speakers who are close and who interact frequently. To account for this interaction, Geoffrey Leech posited a special maxim that he called the **banter principle**.

Sometimes this banter becomes more like a ritual slanging match. In her honours thesis on teenage slang, Wendy Allen gives the following example (from urban Australia). As you can see here, teenagers trying to outdo each other by taunting a mate with insults about them or their family in front of an audience.

A:If I had a pussy like yours I'd take it to the cats' home and have it put down …

B:If I had brains like yours I'd ask for a refund …

A:Well, if I had tits like yours I'd sell them off for basketballs …

Ritual abuse of this nature is a competitive game – a kind of teasing. It utilises the same categories as the kind of insults to outsiders (or people cast as outsiders). As a display of one-upmanship, these interactions use insults based on people's (supposed) sexual practices, age, appearance (body and clothes), smell, domestic arrangements, and so on – precisely the same categories that you find in true insults, intended to wound, humiliate and belittle. However, they aren't an attack on an enemy or someone who is an outsider despised or disparaged, but rather an expression of group solidarity. Normally, abusive address forms or epithets are uttered and reciprocated without animosity – and it all indicates strong bonds of friendship.

Ritual slanging matches are found in many communities, past and present. Late nineteenth-century American cowboys engaged in *cussing contests*, where a saddle would be awarded to the most abusive participant. *The dozens* is the term used of the same behaviour among African Americans today. The dozens is also called *bagging, capping, chopping, cracking, cutting, dissing, hiking, joining, ranking, ribbing, serving, signifying, slipping, sounding* and *snapping*. There are even earlier examples of something called *flyting*; this was around in Anglo-Saxon times and continued into the fifteenth and sixteenth centuries. Hughes, who has studied the evolution of English swearing habits, describes it as 'the fine art of savage insult'. In true banter fashion, flyting contained every conceivable type of insult, including the sexual, religious and scatological – and it could also demonstrate considerable linguistic skill. The language could be sophisticated, passionate, highly charged and very foul – rather like a very early version of the insult-trading rap battle today.

14.8 Gesture in communication

> Gesture is not just for Italians …
>
> Abner, Cooperrider and Goldin-Meadow (2015), 'Gesture for Linguists: A Handy Primer'

Alongside speech and signing, face-to-face communication is generally supported by gestures and, like language, gestures have cultural variants. In this section, we outline some different types of gestures and give culturally distinct examples.

Gestures can be classified in a couple of different ways. One approach is to sort them in relation to how they are made – for example, with the hands, head, lips and so on. Another approach is to examine their communicative function. In this approach, there is a distinction between **interactive gestures** and **representational gestures**. Interactive gestures contribute to talk by allocating turns, raising eyebrows to elicit feedback and so on. Nods of the head that mean 'yes' (or perhaps 'no') and flicks of the hand to bring someone closer or send them away also fall within this broad category.

Representational gestures (which accompany the presentation of an idea – very common in TED Talks, for example) indicate something about objects and events being described by the speaker or signer. For example, chopping vegetables may be accompanied by a sharp vertical downward gesture of the hand or the introduction of a large object may be accompanied by an expansive gesture.

Within the category of representational gestures, we can also distinguish between **deictic gestures**, which point to objects or locations in space; **iconic gestures**, which reflect properties of the object being described verbally, or **conventionalised gestures**, which are the most culturally specific – these include the specific meanings of head nods and thumb and finger gestures in various communities. Although it can be helpful to recognise these different aspects of gestures, it is also important to acknowledge that these different strategies of meaning making: deixis (mentioned in Chapter 13), iconicity, and conventionalisation (both mentioned in Chapter 1) tend to work together in the development of gestures.

A great example of gestures that are different in different cultures is finger counting methods. Depending on where you are, the number four can be indicated by any of the following gestures (see Figures 14.1).

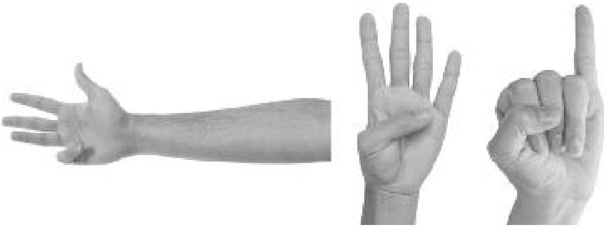


Figure 14.1 Gestures that indicate the number ‘four’

These gestures vary according to whether the extended fingers represent the starting position or the end position, and whether counting commences from the thumb or index finger. The one permutation we haven’t seen is the closing of four fingers with the thumb extended – surely a sign that would have been disrupted by other uses of this hand shape in modern times. This can still mean the number one in systems where hand counting begins with a closed hand and the thumb is the first digit raised (see Figures 14.2).

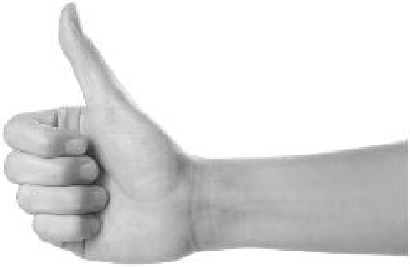


Figure 14.2 A gesture that expresses the number 'one'

Teasing apart sign language and gesture

There is a growing body of research that is investigating sign language and its relationship to gesture. Both sign language and gesture make use of the same visual modality. In each case, communication relies on observation of the speaker's body by other participants. How does the role of gesture in spoken languages compare with the role of gesture in sign languages? Hodge and Johnston explored this topic by contrasting spoken utterances such as the following with similar communications in sign. The following spoken remark, for example, was accompanied by pointing gestures that indexed the referent and supported deictic reference:

> And a corpus file right now, **this** is the first file of the Brown Corpus, and we can see it shows up **there** six times. Okay. One time so in the first example and the first case this is the sentence and it shows up **here**, second one it is here, third one **fourth, fifth** and **sixth**.

In describing signs in languages such as Auslan, it is useful to distinguish between lexical, partly lexical (or partly conventionalised) and non-lexical signs (spontaneous, often indexical or iconic). Partly lexical signs include pointing signs, which identify participants and track them through the discourse, and depicting signs, which are used to communicate properties of the referent. Non-lexical signs rely heavily on context for their interpretation. Hodge and Johnston found that fully lexicalised signs were more typical in slots associated with new information (that is, late in the clause; see also Chapter 13), while partly lexical and non-lexical signs were common in places where old, or established, information was found at the start of clauses.

In the same way that gesture is closely linked to the spoken message, Hodge and Johnston's work confirms the tight integration of the 'linguistic' and the 'gestural' in native Auslan usage – signers exploit gestural elements (alongside grammar) to construct meaning, just as speakers do in their everyday conversation.

Emoji as digital gesture

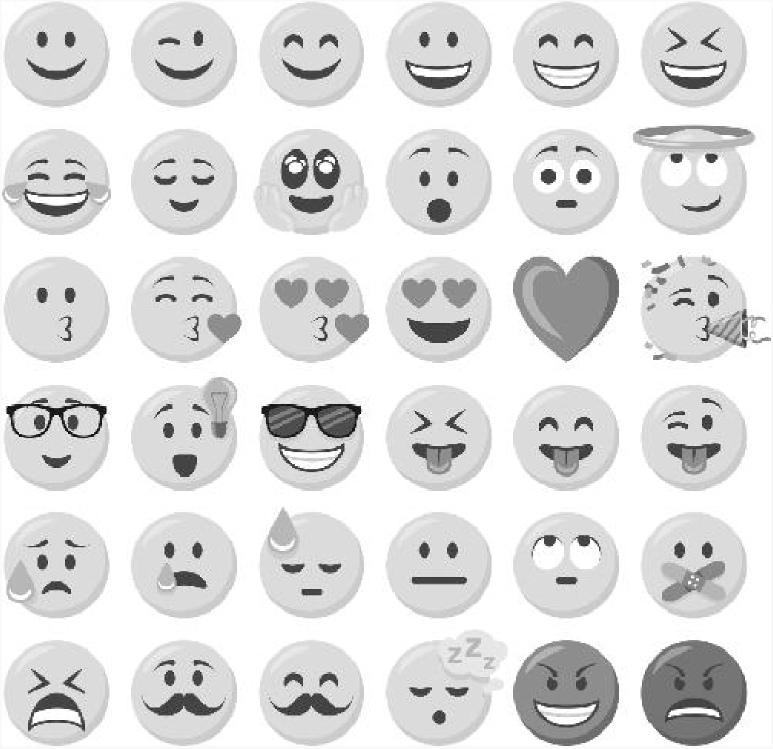


Figure 14.3 Emoji as digital representations of gesture

As you have seen throughout this book, human languages, signed and spoken, have a hierarchical structure. In the case of spoken language, sounds organise themselves into syllables, which combine to form parts of words, which combine to form words, which we can then combine and recombine into an infinite number of different structures. This organisation of level upon level is what distinguishes human communication from that of other animals. People often describe little picture emojis as a language – in 2015, the ‘Face with Tears of Joy’ emoji was even declared the *Oxford Dictionaries*’ ‘Word of the Year’. So, what sort of grammatical architecture might underpin this particular language?

Linguists Gretchen McCulloch and Lauren Gawne have been exploring how people use emoji and punctuation to furnish their writing; as part of this work, they have examined a large corpus of authentic emoji. They conclude from this that the best way to think of emoji is as digital gestures – not as a language proper with its own hierarchical structure. By examining emoji within the theoretical framework of gestural studies, they show that emoji act in relation to written text in the same way gestures act in concert with speech – for example, repeated emoji (like the repeated thumbs up emoji) serve an emphatic function that parallels the use of beat gestures in spoken discourse (where hand shapes move to reflect different aspects of the prosody of the accompanying speech). It is hardly surprising that works of literature translated into emoji (such as Emoji Dick) haven't taken off.

Further reading

On the conduit metaphor, see Reddy (1979). The expression *girlie-man* has an interesting pedigree, described in Battistella (2006); see also Hurst (2014). On speech act theory, see Austin (1962). The quote in 'When it's best not to tell' is located in Urban (1998: 210) and you can read more about Mali approaches to knowledge management in Stebbins and Planigale (2010) (the quotation is on page 151). The quotes and examples from Matsumoto can be found on pages 405 and 412–13 in Matsumoto (1988). The original formation of the cooperative principle was in Grice (1975), but there have been many developments since this time. Relevance theory is discussed by Sperber and Wilson (1995). For more details on thorny issues such as implicatures, interferences and presupposition, we recommend Allan (2001). Our discussion of swearing comes from Allan and Burridge (2010); examples of ritual abuse come from Allen (1987), the discussion of flyting from Hughes (1991) and the *Punch* comic from Stern (1965). An excellent account of impoliteness is Culpepper (2011); see Leech (1983) for the six maxims of politeness. Face is discussed in Goffman (1967), and Brown and Levinson (1987). Barry Blake's (2007) book on linguistic humour has some classic examples of violations of Grice's maxims in jokes (see pages 126–8). The discussion of cross-cultural pragmatics is based on Ogiermann (2009) and Wierzbicka (1991, 1992); see also Eades (2010) for more on pragmatic challenges in the courtroom. Our discussion of silence in Australian Aboriginal communities is based on the work of Mushin and Gardner (2009). Hofstede (1980) provides a useful framework, based in social psychology, for exploring cultural differences across national groups; on conversational styles, see Tannen 1984. The *Australian Journal of Linguistics* 34(2) is entirely devoted to research on sign language in relation to gesture; you can find Hodge and Johnston (2014) in this volume (the quote is from page 264). Our account of emojis and gestures comes from McCulloch and Gawne (2018); for more on gesture studies, see Kendon (2004).

Exercises and discussion points

1 Performatives

(a) The verbs below can be used as performatives in various particular contexts. The circumstances for some verbs are more general than for others. Try to identify a context in which a performative reading is allowed and create an example sentence to illustrate the performative usage.

(i) find

(ii) reject

(iii) swear

(iv) pronounce

(v) declare

(vi) order

(vii) promise

(viii) pledge

(ix) bid

(b) What other conditions must be met in order for each of these verbs to successfully be used as a performative?

2 Being cooperative on the internet

In this chapter, we saw the different ways in which people can disregard Grice’s maxims (for example, irony, sarcasm, hyperbole and metaphor). Consider various (mis)behaviours on the internet and how they illustrate violations of the ground rules for normal interaction. Here are some suggestions, each with *Urban Dictionary* definitions: trolling (‘being a prick on the internet because you can'); spamming ('posting useless crap on forums over and over'); lurking ('spying on people online, while you remain invisible').

Given what you have observed, do you think it is sensible to apply Grice's maxims to virtual speech communities, or do you think interaction is so distinctive that a completely different set of pragmatic principles is needed? What role do you think moderators play in online communication?

3 Social distance, power and politeness

Watch a soap opera of your choice and note how different characters speak to each other. Do you find evidence for a relationship between the different indirect strategies of communicating (for example, hinting instead of demanding)? If so, what different social relationships among the characters do these behaviours reflect? If not, why do you think this would be the case?

4 On names and naming

Think of all the different ways you have for addressing people by name or referring to them when speaking with others. Compare the possibilities of your variety of English with other languages you know about. Now answer the following questions.

(a) To what extent do differences in power impact naming choices in English?

(b) What about social distance?

(c) If names are not a major signal of social distance or disparity in relative power, what kinds of linguistic practices do reflect these differences within relationships?

(d) To what extent is English like other different languages you can explore?

5 Saving face

Gratuitous concurrence is the practice of saying yes, going along with something even if you don't understand or don't agree. Think of times when you have done this or times when someone has responded to you by doing this. Gratuitous concurrence is often used as a face-saving strategy.

(a) Why would face-saving be needed in your examples?

(b) Whose face was threatened? Why?

(c) Did the gratuitous concurrence have the desired effect?

(d) Were there any costs involved in using this strategy?

Research project

Coming to grips with cross-cultural pragmatics

Interview someone from a culture that you know very little about. Ask them about how politeness is managed in their culture. Ask them questions such as the following.

1 Is it polite to use different ways of naming people depending on who they are? Does this happen only when you address them or also when you talk about them to others?

2 What is communicated through gestures as well as (or instead of) words?

3 What things can you ask of family members? How is this different from what you can ask of your friends? Work colleagues? A boss?

4 If someone offered you something (or invited you somewhere), how would you know the offer was sincere? (How) could you say ‘no’ politely?

5 Use the dimensions of social distance, power and ranking impositions in your descriptions of your interviewee’s responses. How do people experience face-threats in this culture? Is it different from yours? If so, how? What kinds of misunderstandings might be associated with these differences?

For more activities, please go to www.cambridge.edu.au/academic/linguistics.

15

Language and social values

◆

> One goes to the potter for pots, but not to the grammarian for words. Language is already there among the people.
>
> Patanjali, second century BCE, *Mahabhasya* (trans.)

15.0 Introduction

So far, we have been quite focused in this book on the ways in which people in communities generate, use and (mostly) inadvertently change language forms. Along the way, our discussions have also provided illustrations of how people attribute values to different words and structures. Below are just some examples we have encountered.

- Standardisation is understood to involve linguistic ‘best practice’ – correctness, precision, purity and elegance are among the perceived qualities of standard languages.
- In some communities, code-switching is a highly valued practice, while in others it is viewed as sloppy or even defective.
- Social evaluation has a dominant role in the spread of a linguistic feature through a speech community.
- For many people, linguistic change is a sign of decline and decay – if dictionary makers and handbook writers acknowledge change, there are howls about declining standards.
- Minority languages in unequal contact situations are often

held in low esteem – the predicted outcome is decline and shift to the dominant language.

- A 'me Tarzan, you Jane' image of pidgins and creoles is held by many who refuse to accept these varieties as real languages.

We now turn our attention to these sorts of 'commonsense' beliefs that people hold about language and its structure, nature and use. In this chapter, we are interested in ideas about language forms, not so much the forms themselves. We look at how language is evaluated: how people assess varieties of language differently and how they assess the speakers of those varieties. We explore why this happens and what the practical implications of these attitudes can be. In other words, we are concerned with **ideologies of language**, and the types of policies and practices that are associated with them. We also consider briefly how language contributes to ideologies about all sorts of issues in society.

The repercussions of linguistic ideologies can be dire for some groups. Popular perceptions about language, especially beliefs about correctness and incorrectness, can appear harmless, but they can also turn into powerful rhetorical sticks to beat others – effective strategies for maintaining social power and short-circuiting serious debate. In Chapter 10, there was even an example where inaccurate folk linguistic views about how people speak put the wrong person behind bars.

Tapping into popular perceptions

Cara Penry Williams has conducted extensive research into language attitudes in Australia, taking quite a different approach to the one outlined in Chapter 10. She argues that, rather than dismissing popular (or folk linguistic) accounts as bad data (because they reflect inexpert, partial and subjective perspectives), we can use such material to better understand the more conscious and publicly negotiated aspects of identity. In this approach, language variation is important because it is used as an index of the identity of particular groups – something people can do consciously and expertly in various ways. Clearly, there is a strong thread of language ideology to be considered here. In her recorded interviews of young people from Melbourne (aged 20 to 25), Penry Williams found that one noteworthy way in which participants expressed their attitudes to variation was through ‘voicing’. Voicing involves participants giving a performance of the sort of speech on which they are commenting, to demonstrate both the language practices being discussed and the social value being attributed to it. For example, in discussing the difference between [æ] and [ɐ:], one of the participants talks about the word dance pronounced [dɐ:ns] as triggering images of people in furs holding glasses of champagne and saying, ‘Shall we dance [dɐ:ns], darling?’ in posh voices; this painted a comprehensive picture of the sort of person that she associates with this form. The participant could identify that what she was constructing was a stereotype (saying, ‘It’s not really how it is but, yeah, that’s what I see in my head’), which provided additional evidence of the ideological associations the vowel distribution had for her. Importantly, Penry Williams notes that in her interviews, these types of examples were used by participants to construct Others from whom they wanted to differentiate themselves – in other words, the identity of the Self was constructed through statements about how the Self differs from Others.

To launch into this discussion of linguistic ideologies, consider a simple illustration of social prejudice and how it can affect communication between different groups. Many of you will be familiar with Dick King-Smith’s book *The Sheep Pig*, although you may know it better as the film *Babe*. In the following short extract (from the movie script), Fly (the sheepdog) is attempting to talk to a flock of sheep. Fly is used to rounding up sheep, but she has never actually spoken to them.

Narrator: Fly decided to speak very slowly, for it was a cold fact of nature that sheep were stupid, and there was nothing that could convince her otherwise.

Fly: Please, someone tell me … what happened this morning?

Narrator: The sheep decided to speak very slowly, for it was a cold fact of nature that wolves were ignorant, and there was nothing that could convince them otherwise.

Sheep: It was Babe … he saved us … the wolves killed Maa … But Babe drew the wolves away!

Fly and the spokes-sheep have never interacted, and yet they ‘know’ things about each other, and assume all sorts of qualities and attributes simply on the basis of ‘common sense’.

15.1 Linguistic prejudice and stereotyping

> Have you ever noticed how, in Hollywood movies, all the villains are played by Brits?
>
> Ben Kingsley in ‘It’s Good to Be Bad’, 2014 advertisement for Jaguar F-Type

Ben Kingsley is right. In Disney animations, the characters who are voiced with British accents (especially posh ones) are overwhelmingly villains; think of Jeremy Irons as the voice of Scar in *The Lion King*. The casting of the voices in all these animated films is underpinned by language ideology.

Chapter 10 described how people use language to locate themselves inside a social space, and it gave you a small taste of the array of social variation that can exist in speech communities. Obviously, how someone speaks is an important source of information about that person’s social characteristics – information that is handy for speakers to know so that they can tailor their behaviour accordingly. For example, it is helpful to know the sorts of things that a person might value, how that person might react to something we mention, and so on.

The habit we have of forming opinions based on the social signals we find in people’s speech is what constitutes linguistic **prejudice**. In fact, we are constantly judging each other on the basis of all manner of aspects of human behaviour, some very trivial: they like football, we like rugby; they eat meat, we don’t; they grill food, we fry; they drink wine, we drink beer; and so forth. Links are made between these aspects of behaviour and personal traits, such as intelligence, integrity, toughness, reliability and social attractiveness. Forms of speech are associated with the features attributed to the groups concerned – so difference in language becomes not just a source of other people’s judgements but is itself open to judgement. As is usually the case with social practices, what one group relishes, another spurns. And when this in-group favouritism leads to others being treated unfairly solely on the basis of their use of language, linguistic prejudice turns to discrimination.

For example, Mali speakers often enjoy speaking very quickly and their songs are also sung at a very high rate of syllables per second. Doing this takes advantage of the many phonological processes in the language that simplify syllables and limit the number of phonetic stops. To Malis, fast talk is an enjoyable sort of language art; however, to their Austronesian-speaking neighbours, speakers of Baining languages like Mali often sound as if they are saying *blurblurblurblurblurb*. And this apparently fluent nonsense (outsiders rarely learn Mali) is readily associated in the minds of some outsiders with the Baining ethnic group's supposed lower intelligence (a proposition that is, of course, racist, culturist and scientifically insupportable). Similarly, speakers of non-standard dialects of English, and people with writing habits that do not reflect the standard (for example, poor spellers), are easy targets for abuse, justified by those who link these linguistic traits with a lack of truth and clarity.

Ve get too soon oldt un too late schmart

A clear example of linguistic stereotyping is experienced by the bilingual English–Pennsylvania German speakers in the Anabaptist communities in North America (such as the Amish and the Old Order Mennonites). In both Pennsylvania and Canada, there exists a popular perception of this group as speaking a form of jumbled English, or what is known colloquially as *Verhoodelt Englisch*. Descriptions abound in the form of jokes and anecdotes – examples like the following even appear in reputable publications:

> Throw Father down the stairs his hat once. It wonders me if it don't gif a storm.
> Becky lives the hill just a little up.
> Yonnie stung his foot with a bee un it ouches him terrible.
> Ve get too soon oldt un too late schmart.

These popular characterisations appear in children's colouring-in books and on tea towels, beer coasters, serviettes and so on. Like most stereotypes, they are stigmatised, and like most stereotypes they don't reflect actual usage. There is surprisingly little in the way of German interference in the English of Pennsylvania German speakers – certainly nothing remotely resembling what you see in these examples.

These clichéd features of *Verhoodelt Englisch* would have characterised the speech of the earliest German speaking immigrants during the seventeenth and eighteenth centuries. During this time, the original diglossic situation resembled that of modern-day Switzerland, with Pennsylvania German as the L-variety and High German as the H-variety – there was simply no need for the Pennsylvania Germans to learn English. The curious thing about stereotyping is its tendency to fossilise. Despite evidence to the contrary, stereotypes can develop into a kind of time-honoured tradition that seems immune to change.

Moreover, it is not simply speech that provides the clues about a speaker. It can be appearance, skin colour, sex and so on. The horse-and-buggy Anabaptist communities in North America, being so highly visible and so eccentric to outsiders (the broader community), will always be susceptible to this sort of stereotyping.

In Chapter 7, we introduced Alexander Mitchell as someone who pioneered the study of Australian English pronunciation. He also had a lot to say on the ability of speech to arouse prejudice in people. Here is an extract from an article that he wrote in *ABC Weekly* (5 September 1942):

> Question a man's pronunciation of a word and you may touch him as nearly as if you doubted his moral integrity. Differences in political opinion are often more readily tolerated than differences in pronunciation. We are prepared to believe that a man who differs from us in politics may still be a quite reasonable person. But many of us go through life in the comfortable faith that any man who speaks differently from the way in which we speak must be a knave or a fop or a chump.

The gender-exclusive language used by Mitchell probably makes many modern readers' toes curl, but it is typical for the time – as are terms such as *knaves, fops* and *chumps*! Importantly, though, the point Mitchell is making here is as valid today as it was in the 1940s. In the minds of many people, there is a very clear connection between what they view as 'bad language' and moral decline.

Stereotyping is the fallout of a basic human desire to organise and categorise – our need to impose some sort of structure on 'the chaotic jumble' (anthropologist Mary Douglas's description) that is the world around us, and render it understandable. It is difficult to stop people from using social cues to draw conclusions in this way, and entrenched negative stereotypes are hard to shift, even in the face of disconfirming evidence (a speaker who contradicts the group stereotype usually becomes 'the exception that proves the rule'!). One answer would be to replace negative group images with positive ones – but how to go about this would be the story for another book.

15.2 The identity functions of language

Given the low esteem in which so many dialects and languages are held in various contexts, it is striking how persistently they are used.

In the face of public denigration, lost job opportunities, lower educational scores and a host of other sanctions, people continue to use the language with which they identify best. People value language not only in terms of what it can do for them, or the cultural status it may confer, but also because of who they are when they are speaking. Professionals in many countries use English in their work, not so much because they have assimilated culturally to an English-speaking society, but because that is part of their professional practice. However, language is also a site of resistance to cultural assimilation because using a minority or heritage language is a means of asserting identities are important to oppressed groups. This is true even when languages are highly endangered; as we saw in Chapter 12, communities show improved wellbeing when individuals and families reconnect with culture and relearn, or at least learn about, their languages.

While it is true that many people are able to learn different languages and different dialects, it is also important not to lose sight of the fact that language is far from being the only marker of identity or the only way in which people categorise each other. The proposition that learning Standard English will solve all the problems experienced by Indigenous people or migrant groups in countries such as Australia and New Zealand fails to acknowledge that racism is about much more than language, and that the effects of racism are justified by, but not caused by, language differences. Negotiating identity, insofar as this is ever possible, can be a difficult journey.

Passing for Japanese

Linguist Yasuko Kanno has conducted research on the processes of second-language acquisition of Japanese in Japan and noted that, at the same time as teachers expressed goodwill towards their non-Japanese students, there were a range of pressures on students to assimilate:

> Coexisting with this openness to and acceptance of diversity, however, are forces of assimilation. Physically few students stand out … In many ways, their physiognomic similarity makes the ‘mainstream’ Japanese identity more accessible to many language minority students at Sugino. Recent studies in the Western context … suggest that physiognomic differences – that is, being a visible minority – limit minority students’ ability to assimilate: no matter how much they wish to be the same as the dominant group, their physical features mark them as different. But, at Sugino, many

language minority students can ‘pass’ as Japanese if they want to.

Kanno observes that the physical homogeneity of students is matched by linguistic homogeneity. Students consistently avoid using their first languages unless a teacher compels them to do so by asking them a direct question about their language. She notes that minority backgrounds are so well hidden that students may not even realise they are from the same country. She also reports that there is a strong age-related variable here. Younger students are much more open about their multilingual and multicultural identities. She suggests that this reflects the fact that older students have figured out that their foreign identities are not valued and prefer to hide them.

There is no doubt about it – some languages and some dialects are valued far more highly than others. But note that this is very different from the assertion that one language or dialect is intrinsically better than others. Given all we have seen about how languages are structured and how speakers turn them to their own purposes, it is clear that any system that ranks certain varieties above others is doing so according to values held by the people doing the evaluation. All varieties have the same potential for complexity and richness of expression, and there are no linguistic grounds for saying one is better than another. Of course, standard varieties have amassed considerable linguistic resources, and this gives them serious social clout – but any non-standard variety subject to the same historical development would be capable of developing these resources.

And how are youse today?

An interesting example of non-standard English that is widely heard in both Australia and New Zealand is the new second-person plural pronoun *youse*. Much hated by English teachers everywhere and often associated with various sorts of undesirable traits, it is worth noticing how and when *youse* is actually used – in this way, we can better understand what people are really intending when they use it.

- Teachers at a public school in Melbourne’s west report that students do not use *youse* when addressing them in class, but will use the pronoun to teachers in informal settings.

- Teachers at a Catholic school in Melbourne's south-east report that they only occasionally hear *youse* in the playground and always object when they hear it.
- A colleague reports being addressed with *youse* at a shop checkout counter.

The geography and type of schools indicates that there is a clear class difference involved here. Melbourne's west and the public school system are less strongly associated with the (upper-) middle classes than the south-east and Catholic or private schools.

It is interesting to notice how all students in this (admittedly small and patchy) sample recognise there is a level of formality, requiring the use of a more standard variety in the classroom that does not extend to the playground. It is also interesting to reflect on how *youse* is a marker of regard, or perhaps solidarity; here it is also used as a singular pronoun, as the shopping example illustrates (unless we go with the theory that the checkout attendant couldn't count, which makes no sense given that they spend all day working the till and giving change!). The low tolerance for *youse* among teachers suggests that the class associations with *youse* and its absence from Standard Australian English far outweigh, in their view, its value as a marker of positive fellow feeling.

As this *youse* example illustrates, there are two interacting dimensions at work here: status and solidarity. Even within the one speech community, it isn't a simple matter of one group having more prestige than another. Varieties like the standard language are associated with power, education and wealth, and these things are, of course, highly valued by many people. As we described in Chapter 10, such varieties have **overt prestige**. But 'bad language' (language that deviates from the standard) can have a prestige all of its own too – a hidden or **covert prestige**. Many studies have shown that non-standard speech forms can have strong connotations of masculinity, and so are used by men to express male solidarity. In doing so, men (and in some contexts, many women) opt for prestige of the covert sort rather than the overt prestige of the high-status group. Attitude studies show that so-called low or powerless speech styles also win out when it comes to more human qualities such as integrity, social attractiveness and friendliness.

We conclude this section with a famous piece of research to illustrate just how strongly these prejudices can influence us. A male psychology lecturer, who could use both a posh Received Pronunciation (RP) accent and a Birmingham (Brummie) accent, presented the same lecture to two groups of seventeen-year-old students, using a different accent for each group. When asked to write down what they thought of the lecturer, the students rated him more intelligent when he used his RP accent (recall it was the same talk); the students also wrote more about him when he used this accent (they seemed to like him more in his RP guise). In sociolinguistic studies generally, UK listeners rate Brummie accents as the least appealing and least intelligent. It might have something to do with the poor public image of Birmingham City, reinforced by the use of the accent for negative roles in movies and in advertising. There is no linguistic basis for these views. In fact, many linguistic features condemned in the regional dialects of heavily industrialised urban centres like Birmingham are the same quaint rustic forms so admired in picturesque and unspoiled rural parts of the English-speaking world. This ability to hold self-contradictory points of view is a very human quality!

15.3 Verbal hygiene: Doctrines of correctness and linguistic purism

> I think it is fair to say that there is one sociolinguistic universal: older speakers always think language is going to the dogs – I have never heard any reports of a community where older speakers say that the way their children/grandchildren speak is admirable for its beauty and elegance.
>
> Miriam Meyerhoff (2016), *Borrowing from Bislama into Nkep (East Santo, Vanuatu)*

Linguistic purism seeks to constrain the conduct of individuals by identifying certain elements in a language as ‘bad’. Typically, these are words and word usages that are believed to threaten the identity of the culture in question – what eighteenth-century grammarians used to refer to as the ‘genius’ of the language. Authenticity is at the heart of linguistic purism and it has two faces: one is the struggle to stop linguistic change and to retain the language in its perceived traditional form; the other is to rid the language of unwanted elements and to protect it from foreign influences. As sociolinguist Deborah Cameron has claimed, though, the prescriptive endeavours of speakers are more complex and diverse than this. She prefers the expression **verbal hygiene** over ‘purism’ (or ‘prescription’) for exactly this reason. According to Cameron, a sense of linguistic values makes verbal hygiene part of every speaker's linguistic competence, as basic to language as vowels and consonants. Verbal hygienists are the people found in language associations formed to promote causes as diverse as Plain English, spelling reform, Esperanto, assertiveness and effective communication – even something as esoteric as the preservation of Old English strong verbs (such as *thrive, throve* and *thriven*) and the abolition of the aberrant apostrophe (*Plea's flush the toilet*). They also enjoy thinking and arguing about words, correcting the writing of others, and looking things up in dictionaries and usage guides. These activities are born of a fascination for language, but also the urge to improve it and clean it up.

Take an example from Australian English and its relationship with a powerful relative, American English. One consequence of the rise of mass media is that native Englishes are much more open to worldwide influences. There is a pervasive American dimension to much of what is global; indeed, it is difficult to draw a clear distinction between globalisation and American cultural imperialism. Given the dominance of the United States, and the loosening of ties between Britain and its former colonies, it is not surprising to find a good deal of linguistic steamrolling going on, and the Americanisation of English has been a hot topic in Australia and New Zealand. There are identifiable influences on teenage slang and, more generally, on teenage culture, but the impact elsewhere on the language is minimal. Nonetheless, many reactions from older folk are hostile. Australians continue to denounce 'this wholesale invasion and exploitation'. Vehement objections are made to 'American infiltration into our lingo' and 'annoying American habits … spreading to Australia' – blame is laid squarely on 'the invidious impact of American TV' and 'the Microsoft spell-checker'. To add to this litany of purported horrors, here are two more examples of complaints that Kate has received (note, the typos are original).

Example 1: Letter from JP (4 September 2008)

> I have just heard your discourse on the Americanisation of English of ABC Wide Bay. I am one of the population who is *very much against* this phenomenon, particularly on the Australian Broadcasting Corporation, because, after all, I am one part of the public who help to pay the ABC announcers wages. The words which particularly annoy me are cere moany and hurri cane. … On several occasions I have written to the ABC Head Office complaining about these words in particular and received written replies regarding their Word pronunciation computer.

> You made reference to many other words which have been integrated from the French or British language in relation to food, but these are accepted words to describe the article. Cere moany and hurri cane are not!
>
> … If the offenders are so enamoured of the American language that they have to inflict these words on the Australian listeners, they should be made redundant, emigrate to the United States of America, and go get paid by the American Broadcasting Commission.

Example 2: Email from IM (1 March 2008)

> I find your reference to the subjunctive being a relic, offensive. It's not just something from the past. People who have been educated or take an interest in our language still use it.
>
> The only reason it isn't used is it that people are ignorant. Grammar hasn't been taught in school for over 30 years and now our language is suffering. It is becoming a sort of Pigeon English: omitting words such at [as] that and which and ending sentences with a preposition for example.
>
> I use the subjunctive and if people think I'm mixing my tenses then that's their problem. The onus is on them to be educated. Just because the government school system has let them down is no excuse. If something isn't possible, then why confuse the issue by speaking as though it is?
>
> People generally seem to be quite happy to let English deteriorate into a kind of abbreviated American juvenile dialect, but I'm not. I'll continue resist incorrect grammar and American English.

Puristic endeavours promote a kind of mental dishonesty, and those who attempt them soon find themselves up to their necks in contradiction. French and British English borrowings into Australian English 'are accepted words to describe the article' – American additions are not (example 1). Nothing seems to calm these critics of American English, even the fact that some beloved Australianisms were originally American (the influx of Americans to the goldfields from the 1850s provided many current favourites in the Australian lexicon, such as *bush* and *squatter*). The email in example 2 concerns the demise of the subjunctive, but IM links this to the overall Americanisation of the language. (Ironically, American English is conservative in this regard and better preserves the ailing English subjunctive.)

It is perhaps the high visibility of spelling that intensifies the widespread perception of American influence. When, in 1969, the

Australian state of Victoria advocated spellings such as *color* and *honor* in place of *colour* and *honour*, writers ignored the edict: ‘why should our spelling be changed to follow the American pattern?’, one writer complained to *The Age* newspaper (9 October 1969). People’s beliefs about their language world are powerful – in 2001, public pressure persuaded *The Age* to return to the *-our* spelling. (As an aside, *The Age* newspaper was using the *-or* spelling from the very beginning of its publication in 1854, and it was never an inheritance from America.) Even though many prestigious British publications, including the London *Times*, various editions of the *English Pronouncing Dictionary* and the *Oxford English Dictionary* promote the *-ize* spelling in words like *legalize*, most Australians reject it outright because it also smacks of a deference to America. Such lay concerns about language usage are not based on genuine linguistic worries, but reflect deeper and more general social judgements. Hostility towards American usage is born of linguistic insecurity in the face of a cultural, political and economic superpower; American English usage poses a threat to authentic ‘downunder English’ and is condemned.

And it’s not just ‘old fogies’ …

The majority of people who go to the trouble of phoning in to radio stations and writing letters to the editor are typically middle aged or older. So, what about younger English speakers today – those who have grown up with variation and change as facts of linguistic life? In 2008, Naomi Ferguson surveyed 93 first-year linguistics undergraduates at Monash University. Her study revealed that these young speakers overwhelmingly showed intolerance towards language change, especially when it came to American English influence. Of the 93 students surveyed, 81 per cent expressed the view that the incorporation of American elements into Australian English was detrimental to the language. The following are typical of the explanations they offered.

- ‘Because Australian English would then slowly perish and it won’t be unique anymore.’
- ‘Americans do not speak to the Australian identity.’
- ‘Often US English seems to use “wrong” words, I don’t like the use of “z” instead of “s” and cannot stand “for free”.’
- ‘Why would we want to speak American English? I think “they” are lazy with language.’

- ‘American accents are so nasal and it sounds yuck. American rap terms ≠ cool.’
- ’The pronunciation is degrading the language, especially among the younger generations.’

In 2009, Melissa Wren carried out a survey of a range of different age groups, focusing on attitudes to spelling. One question specifically asked participants about the acceptability of alternative spellings. The following responses are representative of the mindset of the 20–30 year age group (curiously, Wren observed a greater tolerance of spelling variation among those over 50 years of age).

- ‘I think it’s important … not to fall for America’s kind of childish bastardisation of a very old language.’
- ‘I think it’s bullshit.’
- ‘The “other” spellings of words are usually American – negative language traits.’
- ‘I think most instances of this have come from America and since we are not American this should be avoided at all costs and never enter into school English programmes.’
- ‘British spelling is what Australian spelling is based on, and it should remain that way.’
- ‘Sometimes diversity in language … marks a loss of cultural identity.’

These respondents have all gone through the ‘language in use’ approach at school, and in the case of Ferguson’s students, have also had at least one semester of linguistics and been immersed in the accepted wisdom of the discipline. Yet, there is little evidence of any new open-mindedness in their linguistic thinking when it comes to perceived American influence. Clearly, it is not simply older and linguistically insecure listeners or viewers of Australia’s national broadcaster who feel strongly about the threat posed to their language by American English usage.

It is interesting to note that disparaging commentary around American influence wasn’t apparent until American cultural products had well and truly embedded themselves and become a way of life in Australia. While people often commented on marked differences in

American English, little was made of their influence on the home-grown variety – certainly there was none of the intensive fear-mongering that characterises public opinion today. Monash University researcher Lee Murray has been trawling through newspapers of the 1800s for evidence of early views on Australian English. Here is an extract from one of the letters in her corpus – from A. Richards (of North Fitzroy), writing to the editor of the Melbourne *Argus* (4 January 1894) on the pronunciation he was hearing in the state of Victoria [emphasis ours]:

> Take class for class, the Victorians speak English more correctly, as I said before, than the Canadians, Americans, or the Englishmen born and bred. The Rev. Mark Guy Pearse, when on a visit to these colonies a few years ago, expressed this opinion also. No doubt the Victorians of the next generation or two will speak worse English than they do now, but *I very much doubt if the English tongue will 'Americanise' itself judging from the present.* –Yours, &c.

True, remarks could be critical of Australian English, particularly its accent, and parallels were occasionally drawn between 'familiar atrocities amongst young Australians' and 'the Yankee twang' (to cite from one editorial, 'Australian Pronunciation', *Supplement to the Bendigo Advertiser,* 7 November 1888, p. 2); however, there was no sense that these atrocities were American-made, and there were certainly no fears voiced around any cultural bulldozing that might be putting Australian identity in jeopardy. As documented by historian Joy Damousi, however, everything changed with the arrival of the 'talkies' in the 1920s. This change is certainly borne out by the linguistic commentary in the media at this time:

> It must be already apparent to many thinking people that since the introduction of the American talking films … we are in grave danger of the Americanisation of our speech … I do hope that this matter will be taken up by all those possessing the true Australian spirit, and help to save their country from this wholesale invasion and exploitation by a foreign Power. (*Sydney Morning Herald,* 23 June 1930, p. 5)

Australians continue to denounce 'this wholesale invasion and exploitation'. Nothing much has changed – except that the blame is now laid on to 'the invidious impact of American TV' and 'the Microsoft spell-checker' (to cite again from two recently received emails).

Blue-eyed English

Dictionaries and handbooks provide acts of linguistic purism with a more public arena; however, ideas about ‘better’ and ‘worse’ speech exist in all speech communities. Certainly, the history of English has plenty of examples of cleaning-up activities taking place well before people started to lay down the laws on standards. For example, attempts to expunge loanwords and restore lost English vocabulary started in the Renaissance period, long before the creation of any English language dictionaries as we know them.

Hostile responses to linguistic aliens continued, involving influential writers such as Charles Dickens and Gerald Manley Hopkins right through to the modern-day conlangers (constructed language users). Even Australian-born composer Percy Grainger was obsessed with the ‘purity’ of Anglo-Saxon; his ‘blue-eyed’ English, as he called it, also sought to erase all non-Germanic elements. Over the years, some have tried to revive lost native vocabulary, creating modern versions such as *raintilt* to replace *umbrella; inwit* for *conscience*; and *oversend* for *transmit*. Others, such as the poet William Barnes (1801–86), have even gone as far as coining words – for example, *speechcraft* for *grammar* and *mateword* for *synonym*. How do you feel about *outgates* instead of *exits*?

Verbal hygiene involves a number of different types of restrictive practices, including rules of usage and punctuation (often referred to together as ‘style’), along with prescriptive rules of grammar, ideas about things that cannot be said because they are politically incorrect and arguments about how women are represented in language. All of these are expressions of language ideologies. Even though they start out being about forms of language, all of them are also about values. Each of these topics is discussed in Deborah Cameron’s book *Verbal Hygiene*, and in the remainder of this section we present our own brief examples of these restrictions to draw your awareness to the scope and depth of language ideology in society.

Style guides

Style guides provide writers with useful guidance on matters of writing convention. As we have shown in both Chapters 10 and 14, how writers write has consequences not only for their message, but for the way they will be viewed by readers. Anyone who has missed a spelling error or two in an essay, and been labelled sloppy and heavily marked down will know what we are talking about here!

Some style guides, written by people who are constructing themselves as experts as they write, can at times assume intimidating authoritative tones. Consider the following description of the classic text *The Elements of Style*, in which E.B. White, author of the beloved children's classic *Charlotte's Web*, describes his mentor's attitude to points of usage (White revised and expanded Strunk's original book for republication):

> Professor Strunk was a positive man. His book contains rules of grammar phrased as direct orders. In the main I have not tried to soften his commands, or modify his pronouncements, or to remove the special objects of his scorn. I have tried instead to preserve the flavor of his discontent while slightly enlarging the scope of the discussion.

Not all style guides take quite such an authoritative approach, but even the more measured ones proceed on the basis that there are better and worse ways of writing and it is in the interests of the writer who consults the guide to conform. In this environment, it becomes risky to write in stylistically marked ways. The writer is left wondering whether the reader will take stylistic innovation in good humour or tear the writer down as a result. The extent to which writers are prepared to be creative could quite sensibly depend on how much stylistic variation is conventional in a particular genre.

Good grammar

Most professional linguists at social interactions involving large groups of people they don't know well end up spending some time explaining what linguistics is all about. Often it amounts to something along the lines of being about 'how language works'. Sooner or later, we run across people who respond with, 'Well! I'd better watch what I say!' Often these people also make a prompt departure and spend the rest of the evening on the other side of the room. Expertise in language seems, for some people, to imply a sort of moral superiority that linguists as a rule are not inclined to exercise – most of us are enthusiastic promoters of more descriptive approaches.

Where does this fear of speaking wrongly come from and what does it imply? 'Grammar' becomes a flash point for issues of style and value judgements attributed to speakers of non-standard varieties. A useful point of clarification here relates to the different meanings attributed to the word *grammar*:

1. the study of the way we use language
2. rules of correct usage

3 structural relationships within language that allow words to be combined in meaningful ways.

Within linguistics, grammar is used to refer to the third set of concerns. Out in the wider world, the meaning can be both broader and more sinister – there are contexts in which a lack of ability in the standard language can be seriously disadvantageous. As we will see in a moment, it may even disqualify you from holding rights as a citizen.

Verbal hygiene can sound the death-knell for a language under threat

For many speakers, the structural purity of a language and lexical purity in terms of vocabulary are part and parcel of a language's survival. As linguist Nancy Dorian shows so clearly, a turn towards purism is often felt strongly within language revitalisation movements, with people insisting that the way the language is spoken in the present reflects past conventions and practices without alteration.

This yearning for the past and desire to clearly differentiate the traditional language in relation to dominant languages like English are understandable. Linguistic purism appears to be something that is deeply ingrained in the human psyche. But, puristic attitudes can be a real barrier to natural healthy change – change that all languages need if they are to remain viable and versatile tools for a society. Constant criticism and correction can have unintended consequences if young people, and indeed learners of any age, come to feel that the 'proper' Ma-ori or the 'proper' Irish that is presented as the gold standard is just too difficult to learn. Because they already speak the dominant language, it can be easy to give up on learning a traditional language under these circumstances – particularly in the context of a range of other health, social and economic challenges.

It can be challenging to balance an approach that is accepting of variation associated with language acquisition against the desire to bring high standards to the way the language is taught and used. Sometimes accepting that people are more or less careful in how they speak in different contexts – a strategy widely used for large, standardised languages – can be a useful place to start.

Political correctness

Political correctness is broadly concerned with monitoring language use to ensure that it does not cause offence to others. *Pilot* has radically different connotations in terms of expertise than *aviatrix; African American* emphasises not genetics or appearance but the historical roots of a group that now forms part of the United States (thus bringing the name into line with other ethnic minorities, such as Japanese Americans and Italian Americans). The extent to which we choose to make politically correct choices will depend on our ability to use language effectively and our willingness to comply. Compliance, in turn, may reflect a range of values and power structures. Compliance can be chosen, but in some circumstances – and particularly in relation to the requirements of human rights or sex discrimination legislation – it can also be compelled.

Pronouns with attitude

Changes carried out in the name of political correctness aim to remove the stigma of negative social stereotypes by compelling us to go beyond the simple content of a message and confront the prejudices embodied in language. So, 'PC' language is a form of public action – it deliberately throws down the gauntlet and challenges us.

English has no gender-neutral pronoun set aside for both sexes, although masculine *he* has been promoted to this role by prescriptive grammars. However, the pronoun *he* is not gender-neutral, and never has been, and to make this point some people use *she* instead: 'When the student has finished this book, she will have a good grasp of linguistics'. Now, pronouns are little grammatical words that are supposed to take a back seat in discourse and writers who use *she* in this way are deliberately pushing the pronoun out of its backseat position in order to make a point. 'Generic' *she* is intended to jar – by drawing attention to form, it forces us to sit up and take notice. Even the use of the double pronouns *s/he* and *him/her* can be deliberately clumsy, the offence of a distracting style being an effective way of getting a message across!

Of course, another widely used strategy is the plural pronoun *they*; this pronoun also has other recently developing uses. For intersex people (those born without typical physical or chromosomal features associated with male or female bodies), and for people who simply don’t identify strongly with one gender or another (and who may prefer one of a range of identities including queer, trans or non-binary), English sets up a pronoun problem that is difficult to ignore. In order to avoid the gender assignment required by using a third-person singular pronoun, many people with a non-binary identity prefer to use *they*. Other forms, in complete pronoun sets, have also been tried and the discussion has been going on for quite some time: *hu, per, ve, xe* and *ze* (among others) have been promoted in various places.

Objections to political correctness can be roughly summed up as being about all the things we’re not allowed to say and the ‘bastards’ (or perhaps ‘bitches’) who won’t let us say them! Cameron notes in *Verbal Hygiene* that these tussles (Are women *pilots* or *aviatrixes*? Is a person in a wheelchair *handicapped* or *a person with a disability*?) are less interesting from a linguistic perspective than the fact that political disputes get so tangled up in the use of particular words. Many people dislike the fact that in the political correctness arena, words become political gestures, and any violation of ‘PC’ protocol becomes an inquiry into their character (if they express themselves in an ‘incorrect’ way, they will be judged). Many complain about losing what they see as their freedom to call things by their ‘right names’.

Of course, people don’t like linguistic change either – especially change that smacks of deliberate manipulation (speech codes, guidelines and so on). Hostility to political correctness has also been fuelled by reporting of stories of over-the-top speech codes and book banning (headlines such as ‘It’s a Sexist, Racist, Fatist, Ageist World’, ‘Correct Thinking on Campus’, ‘No Free Speech Please, We’re Students’, ‘Will Political Correctness Kill Free Speech Here Too?’, ‘Gay and Jew Get New Deaf Signs in NZ’). Because it is politically driven, this language will attract more attention, and certainly more negativity, than most linguistic prescriptive practices. Curiously, though, political correctness has been quite successful at getting people to change their linguistic behaviour – more successful than other kinds of prescription, such as punctuation (people generally don’t like being told what to do with their language!). But perhaps this success is simply the fallout of natural linguistic evolution in the face of more general social change – the inevitable outcome of successful equal opportunity practices.

Language and gender

Along with tussles over the way women are (mis)represented in writing, there are other issues that challenge women as a particular group in relation to language. It is widely acknowledged that women and men speak differently, and this is hardly surprising given the importance attached to gender identity. People with ambiguous appearances or who flout **gender conventions** are sometimes subjected to vociferous attacks, and for this reason a lot of women and men conform (not necessarily deliberately or consciously) to gender norms.

A clear example of the gender dimension to verbal hygiene is the different way in which society evaluates male and female talk, in particular the age-old belief that women talk more than men. English literature is filled with characters that support the stereotype of the chattering woman. Rosalind in Shakespeare's *As You Like It* says at one time, 'Do you not know I am a woman? When I think, I must speak.' Proverbs abound in languages from around the world, all expressing a similar view.

- *Spanish:* 'A nightingale will run out of songs before a woman runs out of conversation.'
- *Hebrew:* 'Ten measures of verbosity came down from heaven, says the Talmud, and the woman took nine of them.'
- *Chinese:* 'A woman's sword is her tongue and she does not let it rust.'
- *Japanese:* 'Where there are women and geese there's noise.'

Despite this widespread belief that women are more talkative than men, research findings constantly contradict this. Here's what linguist Jennifer Coates has to say on the subject:

> Men have been shown to talk more than women in settings as diverse as staff meetings, seminars, television panel discussions, mock jury deliberations and experimental pairs … When asked to describe three pictures, male subjects took on average 13 minutes per picture, compared with 3.17 minutes for female subjects.

New Zealand linguist Janet Holmes has found the same pattern. After analysing the discussion time at 100 public seminars, she found in all but seven, men took up more of the airtime than women; furthermore, males asked around two-thirds of the questions.

So, how do we explain the spurious portrait of the garrulous woman? One reason is surely that we have quite different expectations of male and female talk. In many cultures, the silent woman is held up as some sort of ideal type. As Sophocles wrote in the fifth century bce, ‘Silence makes a woman beautiful.’ So if silence is deemed perfection, then any talk will, of course, be too much talk. It also has to do with the type of talk. Oscar Wilde put it this way: ‘Women are a decorative sex. They never have anything to say, but they say it charmingly.’ Women’s talk is often depicted as trivial prattle. The nature of such stereotypes is that they are self-fulfilling prophecies. Part of our stereotyping behaviour is that we take note of those features that confirm our expectations; they overshadow other features and become the main characteristics of that person. Any talking will then be perceived as talkativeness.

But, it is actually more complicated than this. When evaluating talk by gender, we need to take context into account – in particular, a difference between the public speaker and the private speaker. Women and men often use talk for different purposes. Whereas males appear to contribute more to talk in public and formal domains (such as seminars, committee meetings and school classrooms), women are more likely to hold their own in a more relaxed informal environment, and are often more comfortable contributing to private speaking than public speaking. This sort of private talk is not so much concerned with asserting status and power; it is more about developing and supporting social relationships. Janet Holmes puts it neatly: ‘[M]en’s talk tends to be more referential or informative, while women’s talk is more supportive and facilitative.’

Clearly, women aren’t the empty-headed chatterboxes depicted in the folk wisdom of those proverbs with which we began. And nor are men the strong silent types that are so often admired. These are fictional constructions that have no basis in reality; nonetheless, they are powerful stereotypes and we need to be aware of them – especially if it means some conversational participants are getting a raw deal.

‘Painful – almost fingernails on a blackboard’

Julia Gillard served as the 27th Prime Minister of Australia and leader of the Australian Labor Party (during 2010–13); she was the first woman to hold either position. Of course, politicians are often criticised for their language, but Gillard’s language (especially her accent) was subject to intense and bigoted scrutiny during the time she was in office. The fact that she is female seemed to make her accent doubly offensive to many Australians – as sociolinguistic studies show, accents perceived to be coarse or broad are typically more negatively evaluated in females than males (interestingly, snapshots of Gillard’s accent showed her accent to be not strikingly broad at all). Gillard was also denounced for many other linguistic features, such as ending sentences with prepositions, speaking in the passive voice, using subjunctives and too many adverbs – all features widely employed by other politicians, and indeed by English speakers generally.

15.4 Prescription, competence and social inclusion

> Linguistic bigotry is among the last publicly expressible prejudices left to members of the western intelligentsia.
>
> Deborah Cameron (2012), *Verbal Hygiene*

Linguistic intolerance is rampant. People frequently make extremely negative value judgements about others who use vocabulary, grammar and accent that they view as bad English, often castigating such people as ‘uneducated’, even ‘stupid’. The basic human right of respect is understood to mean that people can no longer speak of or to others in terms that are considered insulting and demeaning, yet this behaviour does not seem to extend to the way people talk about the language skills of others. Conscious and unconscious discrimination against speakers of non-standard dialects and low-status accents remains rife, and typically goes unchallenged.

The contrast between prescription and description that we introduced in Chapter 1 is useful here because it allows us to clearly distinguish between language practices that are thought to be good as opposed to the language practices that we can observe. Drawing this distinction allows us to ask a number of important related questions, such as those below.

- Who is determining what counts as good language?
- What criteria are being used to distinguish between good and bad language?
- Whose interests are being served by maintaining these distinctions?
- Whose interests are not being served?

Reflecting on these questions makes it possible to be aware of the social relations involved and the political agendas that support them. Note that we are not advocating a value-free approach to language –

that would be impossible, and would also mean we would lose all the ways that language can be turned to constructive purposes. We are drawing your attention to the ways in which language connects to wider social issues relating to identity, power and status – for good or for ill, and often both at the same time, depending on the particular social group to which you belong.

Speakers of non-standard dialects have comprehensive systems of linguistic conventions to follow, and by and large follow them quite faithfully. All dialects have rules; they just do things differently. Sentences such as *I done all the cookin' meself* or *I don't want nothin' to eat* are not errors of English, but rather differences from Standard English; however, to a chauvinistic speaker of the standard, it may appear as if someone who utters these is an incompetent speaker of the language. Consider the implications of the following quote from the United Kingdom Department of Education and Science, which implies that only speakers of Standard English are able to participate in a democracy. (The reader will, no doubt, notice that this quote also indicates that participation in a democracy amounts to understanding what you are told and acting accordingly – but that is yet another story.)

> People need expertise in language to be able to participate in a democracy. There is no point in having access to information you cannot understand, or having the opportunity to propose policies which you cannot formulate … A democratic society needs people who have the linguistic abilities which enable them to discuss, evaluate and make sense of what they are told, as well as to take effective action on the basis of their understanding.

A rather large leap is made in the above quote between the willingness or ability to produce Standard English, as it is understood by the authors of the report, and the ability to understand materials produced in Standard English (without which people are considered to not be able to participate in the democratic process). It is difficult to understand how the use of a non-standard dialect could be taken to mean that people should be disqualified from participation in democracy, particularly given the treatment of language in the United Nations Universal Declaration of Human Rights:

> Everyone is entitled to all the rights and freedoms set forth in this Declaration, without distinction of any kind, such as race, colour, sex, language, religion, political or other opinion, national or social origin, property, birth or other status. (Universal Declaration of Human Rights, Article 2)

Confusion around this point can be clarified by noting that, within a human rights-based approach to social participation, not only individuals but also the organisations that serve them have obligations. Thus, nations have a duty to accept and accommodate multilingualism and make efforts to provide interpreters and translations where needed, as a matter of routine, in contexts where participation is important.

The plain language crusade

The Plain English movement is a piece of social engineering that attempts to increase the accessibility of legal documents – the intended audience is the ‘average’ person struggling with the complexities of documents issued from the insurance, legal, government and banking sectors. Such documents are complex (which is why they usually lurk unread at the bottom of drawers!), but the plain language movement has shown that the incantations of even insurers and lawyers do not defy translation.

However, it is easy to lose sight of the fact that, for many, the standard language – especially in its written form – is simply another barrier (and an effective barrier too, because of its clout). Educated Australian and New Zealand English speakers have a huge exposure to writing, but many citizens in each nation don’t have access to the standard in the same way, particularly its written form – and here we are not only referring to speakers from a non-English speaking background. Even Plain English documents use grammatical constructions and discourse structures that occur rarely, if at all, in everyday speech, and they are only acquired through considerable reading and explicit instruction. Anyone drafting plain-language documents needs to bear in mind the straightforward differences between spoken and written language we explored in Chapter 13. Speech is primary – as fundamental to human beings as eating or walking upright – but writing is an additional skill and not everyone is in the position to acquire it to the same degree.

Educational implications of non-standard variation

> On one occasion when she had been irritating me over some little question of English grammar I bit her arm right through to the bone, an action which I have never for an instant regretted.
>
> Noel Coward (1937), *Present Indicative*

The relationship between speakers of non-standard dialects and educational institutions is sometimes fraught. Children who speak varieties that are relatively distant from the standard are disadvantaged when it comes to learning to read because the language in books does not match up with the language they bring to school. In other words, at the same time as they are learning to read, they are also being required to learn a new dialect – the standard. This is true, for example, for many speakers of Aboriginal English in Australia. Their speech is distinct from the standard, meaning that the standard language used at school is a new variety to be learnt. This is an extra challenge that the system often fails to acknowledge.

School systems have varied over time in their acknowledgement of the presence of second dialects among their students and the educational implications these dialects have. In the worst-case scenario, speakers of non-standard dialects are characterised as lazy, uncultured or ignorant, and their divergent language practices are viewed almost as moral failings that require disciplined correction. In some contexts, speakers of non-standard dialects are even characterised as intellectually below par and sent off for remediation in large numbers.

Better recognition of the different varieties of the language being brought into the classroom, and a clear, shared acknowledgement that acquisition of the standard is part of the work of becoming educated, go a long way towards diffusing the moral undercurrent associated with non-standard speech and writing. It also provides teachers and students with a more empowering set of possible responses.

Begin with what you know

Allowing non-standard dialects to have a legitimate place within schools can be anxiety-provoking for both teachers and parents. Where non-standard dialects have been devalued or stigmatised, it seems nonsensical to allow them into a place where right and proper ways of being are supposed to be taught. However, it is clear that acquiring literacy in the mother tongue is potentially a powerful way to lay solid educational foundations.

Tok Pisin, the lingua franca of Papua New Guinea and an English-based creole, is often denigrated by outsiders as baby talk (for example, by Australian citizens, or expats, working in prestigious jobs within the country). This seems to be because so many of the words in Tok Pisin are similar to their English sources but show a range of modifications that remind expats of the modifications in children's speech (for more on this topic, see Chapter 17).

In Chapter 2, we showcased work being done by linguists in education, and we mentioned the book *Second Dialect Acquisition* written by linguist Jeff Siegel in 2012. In this book, Siegel highlights the importance of recognising non-standard varieties in the classroom. He gives the example of when Tok Pisin was introduced as a language of instruction in some schools in the 1990s; he then conducted an evaluation to see what effect early learning in Tok Pisin had on students' later acquisition of English:

> [T]hose who learned initial literacy in Tok Pisin actually scored higher, not lower [in English], than those students who learned literacy only in standard English. Furthermore, on the basis of the test scores, the prep-school children showed significantly higher academic achievement in English across time (i.e. in upper grades as well). This study clearly refuted arguments that using the D[ialect]1 in formal education would exacerbate interference and adversely affect students' acquisition of the D[ialect]2.

15.5 Language planning

One response to the troubles created by languages that do not meet the needs or expectations of their users is to develop and implement policies designed to change either the languages themselves or the ways in which people use them. **Language planning**, as this activity is broadly known, involves educators, policy-makers and linguists seeking to change language behaviour by providing resources or codifying systems of rules that are intended to bring about desirable (for some at least) outcomes (so, yes, linguists do sometimes engage in normative activities such as language planning and policy).

Some of the topics already mentioned in this chapter involved some official intervention and policies affecting language use. They include:

- spelling reform
- plain language
- protecting and upholding standards
- style guides
- speech codes and non-discriminatory guidelines
- the focus on oracy and literacy levels in education.

Language revitalisation in Australia

As discussed in Chapter 12, the Aboriginal languages of Australia have long been subjected to assimilationist policies that saw generations of Aboriginal children forcibly removed from their parents and banned from speaking their home languages. English was the compulsory language of education at residential mission schools and the only option in the homes of adoptive families. Aboriginal people were often moved vast distances on and off missions and reserves; these measures had the result of breaking up language groups and forcing people to switch away from their traditional languages. English was also necessary for Aboriginal people trying to gain access to employment and to make their way in Anglo-Australian society.

The consequence of all this was that the Aboriginal languages of many parts of Australia (originally numbering 200–250 distinct languages) have not been passed down for decades. Nevertheless, many communities remember some words, and Aboriginal people in a whole range of communities are now taking steps to revive their languages. For some, records are good and people have a lot of resources (grammatical descriptions and dictionaries, for example) to help them, but in other situations very little is known (a handful of words and perhaps a sentence or two). Another challenge for communities embarking on this type of work is the quality of the records; the haphazard methods of gathering material (often by officers of the colonies) make for difficult handwriting and erratic spelling.

Language revitalisation has also raised issues for linguists working in the area since this type of deliberate reconstruction and extension of a language from (sometimes) very restricted resources is not yet well understood. Finding ways to partner with communities and support their aspirations for language revitalisation provides us with opportunities to learn more about how people relate to and use language to build systems of communication and celebrate their identity.

Language planning: Who decides?

A constant puzzle, particularly for linguists working closely with speakers of small minority languages, is how to produce materials that actually work – in other words, that will be widely used within communities. Both the current authors can report first hand on this. Kate has had the experience of trying to produce an acceptable **orthography** for the community of Pennsylvania German speakers in Canada. She devised what she felt at the time was the linguistically best system possible – one that was German-based but that eradicated the inconsistent aspects of that system (for example, the representation of vowel length, which in German spelling is unsystematic). Yet, the speakers didn't take to this orthography at all. English was their written language – they were used to English symbols, and this way of representing their language looked too strange.

The priorities of the speech community often appear to contradict the logic that linguists bring to language planning. It is important to work with the competing motivations the speech community is trying to manage. It is tempting to think, for example, that a phonemic orthography will be the best way to make a useful writing system for the community, but there are a great many other issues that a community may address through the structure of its writing system – particularly insofar as it is similar to, or different from, that of their neighbours. A focus on phonemes might seem to provide an objective basis for the decision, and the identity issues the community may be trying to manage might be dismissed as overly emotional – but this is highly unlikely to result in an orthography that the community embraces and uses. (It makes good linguistic sense to adopt *-or* spellings in words such as *color* too, but persuading the general pubic of this good linguistic sense is another matter!)

Similarly, printing books in the best-known variety of a language group might seem appropriate if it is too expensive to print books in all the different varieties. However, unless it is clear that the communities concerned would agree to use the books that are produced, the apparent economic benefits may be unfounded. Printing large numbers of books that nobody learns to read is hardly an example of economic prudence.

Orthography development in Papua New Guinea

Catherine Easton spent several years travelling around Papua New Guinea supporting communities with the process of developing orthographies. Although she was able to share with people the tools of linguistics, and of phonological analysis in particular, she feels that her most significant contribution to the process was her ability to create a space in which people could work through language and identity issues for themselves. Orthographies that get used are orthographies that make social and identity sense, not just linguistic sense. As she describes:

> [L]anguage is much more than analysable patterns of sounds and structures and meanings. Not only is language a tool of communication, but it is also an integral part of human culture. It expresses and negotiates relationships of belonging and exclusion, links with the past, connection with the present and dreams for the future. Language development choices are the result of a complex set of socio-historically specific factors which both produce the options and allow for the renegotiation of possibilities. Consequently, the processes of language development involve the negotiation and explicit representation of socio-cultural identity resulting in the production of visual cultural artefacts such as an orthography, spelling rules, written stories, and word lists.

The following processes have been used in language planning to fulfil the perceived needs of language users. These changes happen slowly – especially if they are left to occur ‘naturally’. When they are promoted by governments or language authorities, success depends on the amount of support available to foster skills in using the new forms of the language in the wider community. Given the importance of language to ethnic identity and all the emotion that brings with it, it is hardly surprising that people frequently are reluctant to ‘allow’ their languages to be ‘tampered with’.

Graphisation

Broadly speaking, **graphisation** refers to the development of an orthography (that is, introducing or changing a writing system, as just discussed). It requires a number of different components to be utilised fully, including the system of representation (for example, an alphabet), spelling, punctuation, accentuation and capitalisation. These choices can be political.

Majority rules?

Among the many ethnic groups within the Russian Federation, the Tatars (also known as the Tartars) are one of a number of groups that speak a Turkic language. These languages were originally written in Arabic script; later, Latin orthographies were developed for them, and then, with the increasing influence of Soviet Russia, Cyrillic orthographies were introduced in Turkic countries. Emigrant communities have tended to use the Latin script for Tatar, and in 1997, at the Second World Congress of Tatars in the Tatarstan capital of Kazan, delegates agreed that the Latin script should be used within Tatarstan as well. This move from the Cyrillic alphabet to the Latin alphabet triggered political ructions within the Russian Federation. As Mark Sebba reports:

> No sooner had the Tatarstan laws come into effect than deputies in the Duma began putting forward amendments to the Federal language laws to outlaw the Tatarstan proposals. The rhetoric of the debate made it clear that Tatarstan was being accused of nothing short of disloyalty to the Russian Federation. The pro-Kremlin media called Tatarstan's proposed switch to the Latin alphabet 'a threat to the integrity of Russia' … The chairman of the subcommittee for language policy, ethnic minorities and indigenous peoples, Kaadyr-ool Bicheldei, warned that Tatarstan was showing a dangerous tendency toward rapprochement with Turkey, and other republics in the Russian Federation might follow suit.

Other developments can change an existing writing system – spelling reform is a relatively common example, but surprisingly difficult to implement. We mentioned issues relating to this in English in Chapter 8, suggesting that there are other advantages to the apparently irregular English system, namely that they allow readers to make connections morphologically as well as phonologically (for example ***create*** [kɹiæɪt] versus ***creation*** [kɹiæɪʃən]. The enormous diversity of pronunciation patterns in English presents significant challenges in this area too – a single spell-it-like-it-sounds system would exclude at least as many readers as it welcomed. And, then, there is always the emotion that accompanies any suggested changes. Spelling is not just a matter of representing speech in writing – the letters become potent symbols with considerable cultural and political baggage, and views on spelling (for example, *honour* versus *honor*) can be passionate and ferocious.

To test this out, try some simplified spellings, such as *rime* and *delite*, on friends and see what reactions you get (in fact, these were the original spellings). They might make a whole lot of sense, but for most people they are too strange to the eye. Moreover, writers make use of precisely these sorts of conventions to indicate a lack of education, so they smack of illiteracy. People get attached to spellings, no matter how absurd or irrational, and simplified spelling never gets a foot in the door.

Spelling tussles in Germany

The spelling reform experience in Germany is a good illustration of the difficulty of implementing such changes. The system was officially reformed in 1996, and even though the reforms seemed sensible and were not terribly large in scale, it took ten years of fierce debate before they were officially sanctioned. Basically, there were three camps: one group welcomed the reforms; another group rejected them outright; and the third group wanted them modified in some way. Then there were the editors and publishers who were used to their own opinions and practices, and were reluctant to relinquish these in favour of the new standards. Additional challenges came from German-speaking countries outside of Germany. Countries such as Switzerland and Austria had their own conventions in place. There are still many people unhappy with the reforms and some have even refused to adopt the changes.

Modernisation

Modernisation generally involves the expansion of the lexicon to incorporate terms for modern artefacts, activities and concepts that are otherwise referred to by code-switches or borrowings using the language of the donor culture. To a certain extent, this practice can be motivated by linguistic purism – a desire to prevent the language from being swamped by lexical aliens. When communities seek to revive or revitalise traditional languages, modernisation can be a hot topic.

For example, in relation to language revival in Indigenous communities, people want to respect and explore the past by focusing on the language as it was spoken at the time of first contact. This can push them to feel that modernisation is unfaithful to the true nature of their language. On the other hand, people also want to be able to use the language in their day-to-day, modern lives, feeling that living languages need to have current relevance. This can be experienced as a pressure to modernise, often by developing large numbers of expressions (using the types of morphological processes described in Chapter 4). Both these needs can be met, but when the linguist is an outsider to the community, saying so is not always helpful. It is important that key people within the community have time to reach their own conclusions about what they need to do.

Standardisation

Standardisation refers to the development of a common standard from among a range of dialects. In diglossic situations, the standard form of the language is considered to be the high variety, used in literature, broadcasting and so on. Minority languages without a standard may struggle if the lack of a standard is a barrier to the language community working together. However, most groups find the process of standardisation itself to be extremely fraught. Basically, no one likes having another variety of the language imposed on them as the standard. In countries where a standard has developed, it is usually associated with powerful elites within the community. For example, Standard English was originally a local (prestigious) dialect of the London–Central Midland region – it just happened to be at the right place at the right time (the United Kingdom's political and commercial centre, at the heart of a rich agricultural area and a growing wool trade, and with a handy proximity to Oxford and Cambridge Universities).

In 1966, the Norwegian linguist Einar Haugen identified four stages in the development of a standard language:

1. **selection** of the language variety to be the standard (it may be the language first written down, spoken by a conquering people, identified with the capital, a favoured dialect, and so on)
2. **codification** of the linguistic features of the standard (this requires the production of grammars, dictionaries, textbooks of instruction and other prescriptive literature)
3. **elaboration** (perhaps also modernisation) of lexicon and grammar to allow for a variety of functions (a standard has to deal with a range of ideas and activities, and foreign and new concepts)
4. **acceptance** of the standard by members of the community (success can come through the prestige and authority of the standard).

Labels such as standard and non-standard English imply a clear-cut distinction between varieties of different status. As always, labels like these obscure the reality of flux and variation. For example, a construction like *usedn't* is frequent in the speech and writing of many educated Australians and New Zealanders – yet it still hovers on the border between standard and non-standard. Some speakers claim they say *I usedn't to sing*. Some claim to say *I didn't used to sing*. The rest are fence sitters in relation to which they feel is more acceptable. It's also the case that, until quite recently, both Australia and New Zealand based their standard on the standard variety of England. Some today continue to defer to British norms, as opposed to Australian or New Zealand usage, although the appearance of distinctly New Zealand and Australian dictionaries, as well as style manuals (such as Pam Peters' *The Cambridge Guide to Australian English Usage*), have helped to establish distinctive standards for both countries. These publications symbolise the independence of these new Antipodean varieties of English and demonstrate their ability to articulate their own norms of language rather than referring back to the source variety (British English).

The ideology of the standard

> That the weather clerk really makes the weather probably none but infants believe, but that language is made by compilers of dictionaries and grammars is a conception not confined to the young or ignorant.
>
> Hans von Jagemann (1899), presidential address to the Modern Language Association

There is a sense in which standard languages pass into supernatural realms. (Some linguists even describe them as 'superstandards' to acknowledge their otherworldliness.) They become an ideal that speakers have for their language, and everyday usage never quite rises to the occasion – not even the performances of 'good' speakers and writers. Editors, dictionary makers and handbook writers, who help to establish and maintain this object of worship, become the ones with the specialised knowledge. They possess the shamanic powers to control events, to diagnose and to cure.

Linguists are in a tricky position here. Their descriptive position is often misunderstood as an 'anything goes' ethos that promotes the abuse of standards. In fact, this position is anything but 'anything goes' – what it tries to promote is a responsible and responsive attitude to language. For linguists, language is a natural (even if social) phenomenon – something that evolves and adapts, and can be studied objectively. But, for many in the wider community, language is more like an art form, something to be cherished, revered and preserved. Understandably, this group rejects the neutral position of the linguistics profession; however, to create a work of art like a standard language is to enter into a partnership with natural processes – prescription would soon render the work sterile and useless.

15.6 Language and nation

There is a strong conceptual association between languages and nations. This ideology typically assumes that nationhood reflects a homogeneous set of people speaking a single language. Thus, in Australia, English is understood by many to be a defining characteristic of the nation. What this view fails to acknowledge is both the internal diversity of the English being spoken – including Aboriginal English and other non-standard varieties – and the large number of other languages spoken within the country: Aboriginal languages as well as the languages of immigrant communities from many corners of the globe.

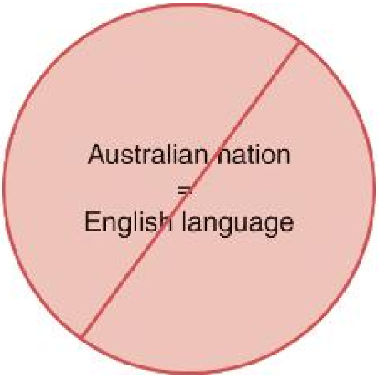


Figure 15.1 Australian nation = English language, an ideology – not a fact

Many immigrants to Australia who speak languages other than English identify themselves as belonging to particular ethnic groups. **Ethnic groups** are associations of people who share historical experiences, some degree of cultural or linguistic similarity and a sense of being different from others.

In Papua New Guinea, the Mali identify themselves as a sub-set of the Baining ethnic group. This group includes speakers of a number of different but related languages who also share a loose collection of cultural practices, quite distinct from their neighbours, along with the belief that they are descended from the original inhabitants of the area. Papua New Guinea is an ethnically and linguistically diverse state. A lingua franca (a common language spoken across different ethnic groups), such as Tok Pisin in Papua New Guinea, has the important function of unifying diverse groups and contributes to political stability.

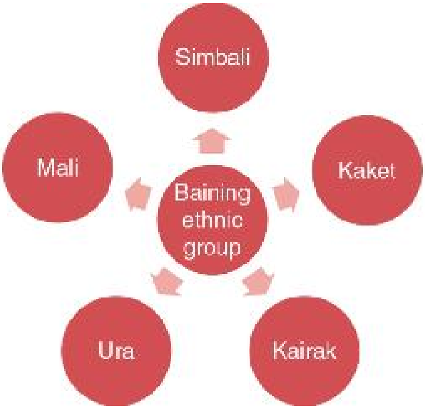


Figure 15.2 The Baining ethnic group

Nations are understood to be associated with self-aware groups that consider themselves entitled to political independence. States are political units covering defined geographical spaces. States can be multinational and nations can involve numbers of ethnic groups. Assertions of nationhood by smaller groups within these larger units are frequently taken to be attempts to gain independence from a larger group. At the same time, this can be a useful strategy for smaller groups that want to be clear about how they see themselves as separate from more dominant groups that surround them. Consider the following quote by Pita Sharples, co-Leader of the Māori Party in 2013:

> So, with pride and humility I stand before you to deliver my address on the State of the Māori Nation. I will discuss issues relating to the scope of the Māori Affairs portfolio, and take time to reflect as Co-Leader of the Māori Party.

This quote drew the following response from the One New Zealand Foundation, a group that objects to any special status being granted to the Māori people. (Note the failure to use the diacritic in 'Māori' here, reflecting a general and deliberate refusal to engage with Māori perspectives.)

> What Maori Nation? Statements like this are a cancer within our society; there has never been a Maori Nation!

No doubt the emotive strength of this response is partly because the writer(s) (who, incidentally, do not identify themselves) recognise that asserting nationhood is a more powerful stance than asserting ethnic identity (a strategy they also contest within their website).

Ideologies of national language planning

The association between language and state can be problematic because it tends to make speakers of minority languages into problems for the nations in which they are found. Reality rarely matches the 'one language, one nation' concept in practice. Language planning often involves adjustments between the language of the state and the language of communities residing within state boundaries. For this reason, language planners have found it useful to tease apart these concepts and hold them in mind as ideas or interests that need to be managed in relation to each other.

The ideas that motivate language planning vary in terms of their focus on accommodating or minimising differences between groups of speakers, but they are not necessarily mutually exclusive. In practice, language planning at the national level involves the following types of policies.

- **Language assimilation** refers to policies designed to enforce the ideal of 'one nation, one language'.
- **Linguistic pluralism** is reflected in policies acknowledging that a range of different languages are being used.
- **Vernacularisation** involves gradual transition to a local language as the official language.
- **Internationalism** increases interaction with others at the international level by choosing an official language that will facilitate communication with the wider world.

'Speak English, you bitches!'

Some time ago in an earlier iteration of the ongoing conflict in the Middle East, a student of ours reported that she was in a Melbourne shopping centre speaking Turkish with a friend who was wearing a headscarf. As the two of them strolled along, a woman with a heavily laden shopping trolley hissed at them, 'Speak English, you bitches!'

The student was in the final year of her Bachelor of Arts with a major in linguistics and worked part-time in a demanding job for a local MP. Both she and her friend were perfectly competent speakers of English, but they were enjoying speaking Turkish together because that was the language they always used when conversing with each other; among other things, it reflected their families' shared histories of migration. There was no separatist intention behind their private, personal choice any more than there was a practical restriction on their speaking English to each other.

The woman who verbally attacked them clearly considered that they were not entitled to make a personal choice about language use in a public setting. When this type of attitude is aired in the press, it is often associated with the assumption that people who speak community languages in public do not know how to speak English and, as such, are failing to adapt to Australian culture. Objections to the public use of languages other than English reflect long-standing assimilationist attitudes and the monolingual mindset we described in Chapter 12.

15.7 Research on language ideologies in practice

> The moment our discourse rises above the ground line of familiar facts, and is inflamed with passion or exalted by thought, it clothes itself with images.
>
> Ralph Waldo Emerson (1836), 'Language', in *Nature*

As we noted briefly in Chapter 13, discourse analysis is the general term given to research about the structure and the effects of language in use. It is concerned with the way that language-in-use interacts with the social, cultural, and political lives and worlds of its users. This large set of research practices ranges from the examples of narrative and conversation analysis we introduced in Chapter 13 to investigations into the ways in which language is used to create knowledge and other types of social categories. Through talk and writing, we construct ourselves and each other not only as (im)polite and (un)cooperative, but (in)sane, (un)emancipated, (un)respectable, (il)logical and so on. Just where we sit on any of these scales depends greatly on the social values and conventions with which we identify, and that we emulate and expect from others. And it doesn't all go one way: where we sit is also an evaluation that others make of us.

For example, continuing the theme of masculinity and femininity we began to explore in Chapter 14 and earlier in this chapter, researchers have been interested in the ways in which language choices both support and construct our ideas about what it means to be male, female, straight or gay. These ideas intermesh with our ideas about politeness (What is a polite woman like? Is she different from a polite man?), but extend beyond them by noting how others' perceptions are important to the results. Judith Butler notes that 'masculine' and 'feminine' genders are separate from our biological sex (which makes sense because there are a range of other combinations besides the XX and XY chromosomes that make up female and male human biology). She argues that these traits are in fact constructed out of things we do, including our language choices, and things others do to or with us, including how we are spoken about. In other words, our gender is maintained through the ways we perform or act it out in relation to social norms. How we use language is one resource for the performance of gender.

Gayness and gossip and straight-talking men

The relationship between language-in-use and the lives and worlds of users is apparent in the types of sanctions that are sometimes (in some settings, ferociously and often) brought to bear on people who transgress gender norms. Research by Deborah Cameron on gossip among young heterosexual male adults explores how complex relationships between language and gender can be. She shows how it is precisely when they engage in gossip about their less-masculine classmates that they themselves reproduce speech with all the features of gossip (overlaps, and collaborative speech, for example). In doing so, they come closest to reproducing styles of language commonly attributed to women:

> In this context – a private conversation among male friends – it could be argued that to gossip, either about your sexual exploits with women or about the repulsiveness of gay men (these speakers do both), is not just one way, but the most appropriate way to display heterosexual masculinity. In another context (in public, or with a larger and less close-knit group of men), the same objective might be pursued through … strategies, such as yelling abuse at women or gays in the street, or exchanging sexist and homophobic jokes. Both strategies could be said to do performative gender work: in terms of what they do for the speakers involved, one is not more 'masculine'

> than the other, they simply belong to different settings in which heterosexual masculinity may (or must) be put on display.

What they talk about, how they talk and the things they say all contribute to the ways in which these young men construct themselves for each other. In spite of the gossipy tone of the interaction, they are aligning themselves against other males in their class who diverge from their ideals of masculinity in all sorts of ways, mostly with no awareness of the sexual preferences of their targets. The confusions within the group about sexuality and gender are highlighted by the complaint from one participant that four *homos* in one of his classes were continually *hitting on* a female student. Cameron suggests that since it is gender rather than sexuality under the spotlight here, this seems to make sense within the group.

Ways of being women and men are valued differently in different settings, as are ways of speaking and seeing the world. Research that began by considering how we make sense to each other (pragmatics) has provided inspiration and conceptual resources for a much larger set of issues relating to how we also make sense of our world (discourse analysis).

15.8 Final thoughts

As we have moved through the book from thinking about the structure of language to thinking about how language varies between speakers and ultimately changes, and on to thinking about how we use language to communicate, we shift our gaze from empirical to more interpretive ways of knowing. Although discourse analysis is a long way from instrumental phonetics in its methods and approaches to data analysis, there are nevertheless key connections between the two, and it is the ways in which social value is attributed to linguistic forms that allow us to navigate these connections. How and why we use language in the ways we do both reflects and creates the language forms, values and routines of tomorrow.

Further reading

Bell (2014) has an excellent chapter on valuing language; an earlier essay on language ideology (available online) is Woolard and Schieffelin (1994). A.G. Mitchell wrote two articles in 1942 on Australian speech in the *ABC Weekly*; these are described in Damousi (2010), a wonderful account of Australian English from 1840 to 1940. The quotation from *The Age* newspaper came from Jernudd (1989). The two linguistic surveys mentioned were carried out by our students, one focusing on attitudes to spelling and the other on student impressions of American English influence (Wren 2009 and Ferguson 2008, respectively). For the classic style guide, *The Elements of Style*, see Strunk and White (1979). For more on the stereotyping and social evaluation of speech, see Giles and Powesland (1975). For the most recent account of verbal hygiene, see Cameron (2012); see also Dorian (1994). For Mary Douglas's anthropological insights into pollution and taboo see Douglas (1966) (the quotation can be found on page 193). Coates (2014) and Holmes (1995) examine the different expectations that society has of male and female talk. Musgrave and Burridge (2010) present acoustic data to assess the extent to which Julia Gillard changed her vowel system after her election to office; see also commentary by Wilson (2010) and Gawne (2011). Siegel (2012), on second-dialect acquisition, was the source for the discussion of educational implications; his quote in relation to Tok Pisin appears on page 240. Our discussion of standardisation was based on Haugen (1966), and Milroy and Milroy (1985); the journal *English Today* (vol. 26, no. 2, 2010) is entirely devoted to prescriptive issues. Although we have not referred directly to his work in this chapter, Bourdieu (1991) provides fertile ground for reflecting on the relationships between language and power; see also Pavlenko and Blackledge (2004). Sharples' 2013 speech is available online, as is the reply by Anonymous (n.d.). The quote about Gurindji multilingualism is from Meakins (2008: 247). Other quotes are from Kanno (2004: 331, 334), Easton (2008: iii) and Sebba (2006: 107). Butler's work on gender is introduced in Butler (1999). See also Cameron (1999) for an application of some of these ideas using the methods of conversation analysis (the quote is from p. 430).

Exercises and discussion points

1 Exploring identity – the culture star

How does language fit with your identity? How many different languages, dialects and registers do you use? When, where and who with? This activity will help you to explore the relationship between language and identity.

Think of all the different groups of which you can be counted as a member and add each one to the points of a star.Here's a star (see Figure 15.3) that works pretty well for either of the authors:

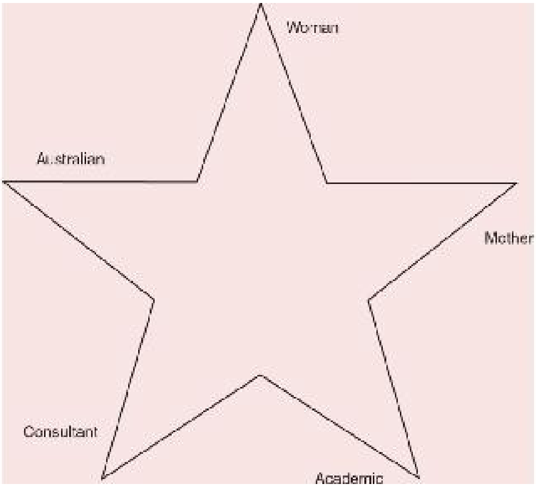


Figure 15.3 An example of the culture star

Even though this star seems to suggest that the authors are somehow 'the same', we defy anyone who knows us to say this to our faces. One reason is that there are many things we do not share that are hidden by this exercise. These include love of the hurdy-gurdy, Western Australian origins, past residence in Papua New Guinea and enjoyment of weeding the garden. Even as women, mothers and academics, our personal histories indicate that constructing us as the same would be extremely short-sighted.

With these reservations in mind, use your star to reflect on or discuss the different sorts of language practices associated with your membership of each group. Are some ways of speaking more prestigious than others? Are some more enjoyable than others? How do these different ways of speaking interact with each other? Do they ever overlap or leak across group boundaries?

2 Changing how you talk – when and why

As participants in a range of social settings, people become adept at changing the way they speak to suit the situation. Focusing now just on how you pronounce words as you speak, think about the following questions.

(a) Are you ever aware of changing your accent or other features of the way you speak (loudly, slowly)? What changes do you make and why?

(b) Do you think that accents should be a topic for teaching at school? Why? Do the identities of the students matter?

(c) Students whose ways of speaking are criticised in the classroom often disengage from education. On the other hand, students may benefit from being able to use a range of ways of speaking. How do you think these issues should be addressed in education?

(d) Is there an English accent that you find especially attractive? Using appropriate terms from phonetics and phonology where possible, describe what it is like.

3 What are the values espoused by this chapter?

In this chapter, we have argued that there is no value-neutral approach to language planning or the promotion of language standards. In his article 'Reply to James Milroy', Cameron argues that linguists 'need to start talking more concretely about what we are for as well as what we are against'.

(a) What values do you think we have promoted in this chapter?

(b) What evidence can you provide to support your findings?

(c) Do you agree with these values? Why? Why not?

4 Feedback on writing

Choose an essay or assignment that you have written and been marked on. It needs to be something on which you have received written feedback.

(a) How much of the feedback is focused on issues of style (word choice, punctuation), spelling or grammar?

(b) Can you identify any parts of your essay that a highly critical reader (imagine Professor Strunk as he is characterised by E.B. White) could additionally have criticised?

(c) Overall, to what extent do you think your mark reflects on presentational issues as opposed to the content of what you were trying to say?

(d) If you had an opportunity to revise, what would you change?

(e) Is there anything you would leave even though it is not stylistically ideal or technically correct? Why?

5 Knowing grammar

(a) How much would you say you know about English grammar?

(b) How have you learnt about grammar and other language-related topics?

(c) Is language, and particularly grammar, something in which you are interested?

(d) Do you believe knowledge of this material is useful to people in general?

Research project

Making sense of verbal hygiene

In the 2012 edition of her book *Verbal Hygiene*, Deborah Cameron proposes a small number of principles that can help to ensure that discussions of verbal hygiene are reasonably constructive. Summarise the principles she identifies (pp. 224–8) and then apply them to a current linguistic controversy.

For more activities, please go to www.cambridge.edu.au/academic/linguistics.

Part 6

◆

‘Language is the dress of thought’

Language, mind and world

16

Language, the mind and the brain

◆

> Though the faculty of language may be congenital, all languages are traditional. The words in which we think are channels of thought which we have not dug ourselves, but which we found ready-made for us.
>
> Max Müller (1873), *Lectures on Mr Darwin's Philosophy of Language*

16.0 Introduction

This chapter opens with a reflection on the relationship between language and the mind. When learning a second language, it doesn't take long for most people to notice that the world is not constructed in precisely the same way in all languages. How does language reflect our understanding of the world? Where do these differences come from? How do they relate to culture? What do they tell us about language and about its human users? Thinking about language and the mind focuses us on our experiences and behaviours as speakers of particular languages. But, the mind is elusive; we cannot answer some of our most basic questions about how language works with a focus on the mind alone.

Shifting our focus to language and the brain allows for quite different sorts of questions. How is language stored in our brains? What thinking processes are involved in producing or understanding speech? How can we find out more about the way language is organised in our brains? These questions become especially compelling when we reflect that brain tissue has roughly the same consistency as cold porridge, is around 77 per cent water and corresponds to roughly 2 per cent of our total body weight. Furthermore, the structure of the brain and its relation to language is only now becoming clearer to researchers, as new technologies and methods make it possible to develop more and more detailed records of the brain activity associated with language.

In this chapter, we will review what is known about how language maps onto different areas of the brain, explore a range of techniques for learning more about how the brain works, and discuss different approaches to understanding how language is used. Here we will draw a preliminary distinction between research that is concerned with representations of language in the mind (how we organise what we know about language) and the study of processes associated with language use in the brain (how we produce and understand language). We close by exploring some of the diversity associated with language and its relation to the brain.

16.1 Language, culture and the mind

> The limits of my language mean the limits of my world.
>
> Ludwig Wittgenstein (1921), *Tractatus Logico-Philosophicus* (trans.)

One of the really fascinating things about languages is the fact that different languages portray the world in different terms. Even things as fundamental as body parts are represented quite differently in different languages. For example, the arm, hand and fingers are all referred to with the single word *angēthik* in Mali, while in Gurindji the lower arm, hand and fingers are known as *wartan*, in contrast to the upper arm, which is called *murlku*.

Kinship terms are a widely recognised area of cultural difference reflected in different languages. Consider the following sets of kinship terms, in Tables 16.1, 16.2, 16.3 and 16.4, from our case-study languages and notice how relations are organised differently in each case.

Table 16.1 Selected Gurindji kin terms

ngaji	father (or his brothers)
ngamayi	mother (or her sisters)
ngapa, papa	older brother
kurrurlu, kapuku	sister, mother's sister's daughter
kalaj	younger sibling
kaku	father's father, his siblings; man's son's child
ngapuju	father's mother, her siblings; woman's son's child
jawiji	mother's father, his siblings; man's daughter's child
jaju	mother's mother, her siblings; woman's daughter's child

Table 16.2 Selected Japanese kin terms

Own family	**Other family**

otōsan	father
okāsan	mother
oniisan	older brother
onēsan	older sister
otōto	younger brother
imōto	younger sister
ojiisan	grandfather
obāsan	grandmother

Table 16.3 Selected Mali kin terms

naunēk	father
nutēk	mother
natka	older brother
natki	older sister
likka	younger brother
likki	younger sister
aik	grandfather
auk	grandmother

Table 16.4 Selected Māori kin terms

matua	father
whaea	mother
tuakana	older same sex sibling
teina	younger same sex sibling
tungāne	brother (opposite sex sibling, female speaker)
tuahine	sister (opposite sex sibling, male speaker)
tipuna	grandfather or grandmother

Comparisons of semantic domains like kinship focus our attention on the ways in which languages differ from one another. These types of differences often prompt questions about how cultural variation – so evident in language – reflects or influences differences in how people think. Indeed, the relationship between language (its rules or structure) and our thinking is something that has fascinated people for centuries. German philosophers of the nineteenth century, like Wilhelm von Humboldt (1767–1835), came up with the term *Weltanschauung*, meaning ‘world view’, to capture this complicated link between language, culture and thought; the word *weltanschauung* has now come into English.

The linguist most associated with this area is Benjamin Lee Whorf (1897–1941). Whorf came to linguistics from a background of assessing claims for fire insurance. What initially attracted him to issues about language were petrol drums. He noted that workers were scrupulously careful with drums that were full of petrol but took an extremely casual attitude towards empty petrol drums. This isn’t a rational way to act, given that throwing a match into a full petrol drum will cause the petrol to simply burn, but throwing it into an empty drum will cause the petrol vapours to explode. Whorf surmised that there was something about the word *empty* that triggered the illogical behaviour of these workmen. Intrigued by all this, he took up the study of language, in particular North American Indian languages, with structures and vocabularies strikingly different from English. His studies led him to the idea that language actually works as a kind of filter on reality; in other words, people who speak different languages have different world-views. Whorf was very influenced by his teacher, American anthropologist and linguist Edward Sapir (1884–1939); hence this idea is sometimes called the **Sapir–Whorf hypothesis**.

‘If Aristotle had spoken Chinese or Dakota …’

Here is an example of the sorts of structural differences that so impressed Benjamin Whorf. It was actually given originally by Whorf ’s teacher, Edward Sapir, and it illustrates the different ways languages express the very simple impression of a stone falling.

In English, we have to specify whether we are talking about one stone or several. We have to state when this event took place and also whether we have a specific stone in mind or just any old stone. In other words, we would typically say, ‘The stone falls’. To us, this seems a pretty reasonable way to report this event. Not so to Russian speakers, however; they would find our attention to definiteness curious – ‘Stone falls’ is good enough. But, not good enough for German or French speakers, who have to assign the stone a gender category (in German, the stone is masculine and in French, feminine). Interesting things start happening to stones when they appear in the Indigenous languages of North America. Speakers of Chippewa (of the Algonquian language family) would also to include the fact that the stone is inanimate, but Kwakiutl (or Kwak’wala) speakers in Canada would go further and specify whether the stone is visible or invisible to them at the moment of speaking and where it is located (relative to the speaker, the hearer or some third party). It’s all mind-boggling detail for speakers of English (or of any Indo-European language), but remember that we have to state that the stone is singular, while Kwakiutl speakers don’t. Their version could apply equally to one or several stones. And they also don’t need to specify the time of the fall either, which we do via a marker on the verb.

So far, the event of the stone falling has been analysed by all speakers as involving two concrete notions: that of the stone and that of the action of falling. This is not how Nootka (or Nuu-chah-nulth) speakers would analyse it, however. They would capture the whole event in one verb form, but nothing like our verb 'fall'. The Nootka verb consists of two main elements, the first showing the movement or position of some sort of stone-like object, the second referring to a downwards direction. Sapir gives the example of a hypothetical English verb, 'to stone down'. Our sentence, 'The stone falls', would be rendered, then, something like 'It-stones-down' (but this is all one word in Nootka, not three as in English).

Linguists have since identified two versions of the proposition that language does influence the way we think. The weaker version of the hypothesis is known as the principle of **linguistic relativity**. This version states that language differences – specifically lexical and grammatical differences between languages – are associated with differences in the way speakers of these languages think. The stronger version of the hypothesis, the principle of **linguistic determinism**, goes much further and asserts that differences within languages actually cause differences in the way people think.

There are lots of different ways to tackle the question of exactly how language relates to thought, ranging from impressionistic and often chauvinistic 'observations' about speakers of other languages (along the lines of French being a seductive language and hence speakers of French being intrinsically romantic types), to claims that because a language does not encode tense (past, present and future), its speakers must not experience time as being linear.

It is unlikely that anyone ever claimed that speakers of Māori cannot distinguish the gender of each of their *tipuna*, for example, but it is occasionally more difficult for people to avoid the trap of wondering whether a Gurindji child can really know which of their *ngaji* claims paternity. Speakers of Gurindji know the second suggestion – that children don't know who their biological fathers are – is just as preposterous as the first – that Māori speakers don't know how to distinguish gender among the over-fifties. Similarly, although there is technically no future tense in English (there's no separate verb form, as in some other languages), speakers of English are fully able to think and talk about the future, as evidenced by all the self-help books that encourage us to stop worrying about it so much! Of course, when we want to express upcoming events, we can draw on a number of different constructions that express different aspects of future time

(*I'm going to/am about to/will/might*, and so on *leave tomorrow*). And, on the matter of 'futured' versus 'futureless' languages, some of you might recall the claim made by behavioural economist Keith Chen that our economic behaviour is linked to whether or not our language has a future tense distinction (check out his TED Talk, 'Could your language affect your ability to save money', and some of the critical reviews that followed). It can be very tricky indeed to navigate a path from the clear indications of how language and culture interact through to an understanding of the relationship between language and thought, while avoiding the perils of over-simplification.

As another example, take those gender (or noun class) systems to which we introduced you in Chapter 5. Recall that some of our case-study languages – for example, Russian, German and Mali – are quite complicated in this area of their grammar, as was English in earlier times (in Old English, all nouns fell into one of three genders or classes: *se hund* 'the dog' was masculine, *sēo giefu* 'the gift' was feminine and *þœt word* 'word' was neuter). So, does any of this grammatical gender stuff affect how we think and how we perceive the world? Most linguists would say 'absolutely not' – grammatical gender is a matter of form and totally irrelevant to the notion of a gendered language. However, in his book *Through the Looking Glass: Why the World Looks Different in Other Languages*, Guy Deutscher describes experiments showing that speakers of languages with grammatical gender do in fact, at a subconscious level, associate properties with non-living objects around them. He describes experiments in which Spanish speakers judged objects such as bridges, clocks and violins ('he' in Spanish) as big, strong or dangerous; German speakers, on the other hand, tended to think of these objects ('she' in German) as having more 'feminine properties', assigning descriptions such as fragile, slender or elegant. As Deutscher describes, this effect was reversed for objects such as mountains or chairs, which are feminine in Spanish but masculine in German. There were also interesting results in another experiment involving French and Spanish speakers who were required to assign human voices to objects in a cartoon. French speakers imagined the 'feminine' fork (*la fourchette*) as speaking in a woman's voice, but Spanish speakers imagined male voices for their 'masculine' *el tenedor*.

It is riveting to contemplate the extent to which these sorts of differences in grammatical structure might actually predispose speakers to a particular line of thinking. It seems commonsensical that our languages do guide our attention in different directions – yet, the problem has always been how to measure this. Empirical testing in the area of grammatical gender is certainly suggestive that speakers of different languages do think and behave differently, and Guy

Deutscher's book concludes that future science (advances, say, in neural networking) will be able to shed more light on the society, language and mind liaison. Experimental tools might eventually be able to show us that this kind of obligatory marking on nouns really does affect the attention and memory patterns of speakers – even shaping tastes and fashions.

Another area of endeavour in this research, which has had a long history, is exploring the way languages describe colours and how colour terms relate to our perceptions of colour. Colour has the advantage of being something that we can perceive without explicit recourse to language, as well as being a set of categories and values that can be named using non-linguistic strategies, such as RGB colour code numbers. Recent research has demonstrated that language-specific colour terminology has a direct influence on the way people perceive colour.

Researching Russian blues

In order to explore how language relates to colour perception, researchers have created all sorts of ingenious experiments that ask people to arrange colours in relation to each other, or to assign colours to existing named categories (for example, by asking an English speaker to assign a turquoise paint sample to either the category of green or blue). These experiments struggle to climb outside of language far enough to really show whether colour perception without overt association with colour terms varies in relation to the way individual languages name colours. The Russian blues experiment, reported by Winawer and fellow researchers in 2007, finally succeeded in demonstrating that language does have an impact on colour perception, even when colour names are not invoked. The researchers used Russian and English speakers and compared their judgements about colours in the range of blue.

This is interesting because there are two basic colour terms for blue in Russian: *goluboy* 'light blue' and *siniy* 'dark blue', but no general term equivalent to the English word *blue*. The researchers wanted to know whether there was a difference in the way Russian speakers perceived a range of shades of blue compared to their English-speaking counterparts and, further, whether the language functions in the brain were implicated at all in the results.

Researchers asked participants to match squares of different shades of blue with a control colour – they found that Russian speakers were quicker to discriminate between shades of blue correctly when the choices involved one from the category of *goluboy* ‘light blue’ and another from the category of *siniy* ‘dark blue’. They also provided evidence that this advantage in processing time related to activation of language in the brain (even though the task itself was not in any way a verbal task) by finding that if participants were required to do another verbal task at the same time (remembering a series of numbers), then the advantage in discriminating colours was lost. In other words, if the language areas of the brain were otherwise occupied, they were not available to support colour discrimination and performance on the task suffered.

16.2 How language maps onto the brain

Tackling questions of language, mind and culture with any level of objectivity inevitably leads to a need to understand much more about language and the brain. Early studies of brain structure, or **neurophysiology**, established that there are localised areas of the brain with specific language functions. Let’s start by looking at the different structures within the brain, then map what we know about language function onto them. Figure 16.1 shows the basic structure of the brain. Many of the structures contained in the human brain are shared with other vertebrates (that is, animals with backbones – everything from frogs to ferrets). These areas include the medulla oblongata, pons, cerebellum, thalamus and hypothalamus (all of which are visible in the diagram), as well as many others. More important for us here is the area called the cerebrum. This large area covers most of the outside area of the brain and is characterised by numerous folds that maximise the volume of tissue that can be held within the available space. The cerebral cortex represents 77 per cent of the volume of the human brain. In contrast, it represents only 31 per cent of a rat’s brain.

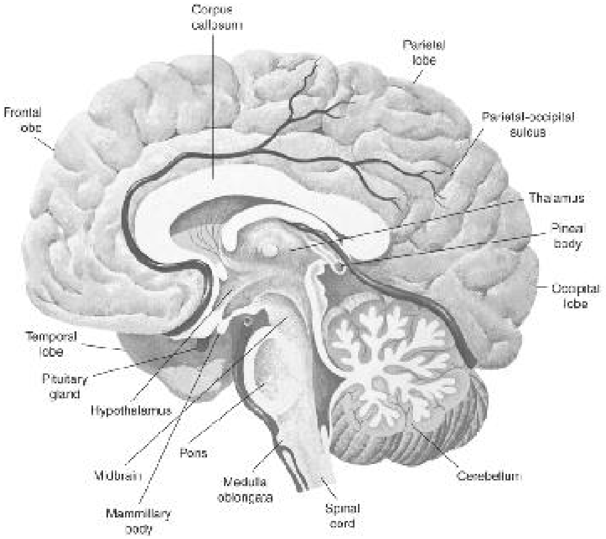


Figure 16.1 Structural organisation of the brain

Figure 16.2 shows how different cognitive functions are localised in different areas of the cerebrum. Of particular relevance to language are the areas associated with speech production, language comprehension, vision, memory and auditory functioning. **Broca's area** is associated with speech production; during the 1860s, neurologist Paul Broca identified the area by conducting autopsies on the brains of a number of patients who had shown evidence of difficulties in speech production while they were alive. Another important area for language, **Wernicke's area**, is associated with language comprehension; the name refers to another nineteenth-century neurologist, Carl Wernicke, who conducted research on the effects of brain disease on language.

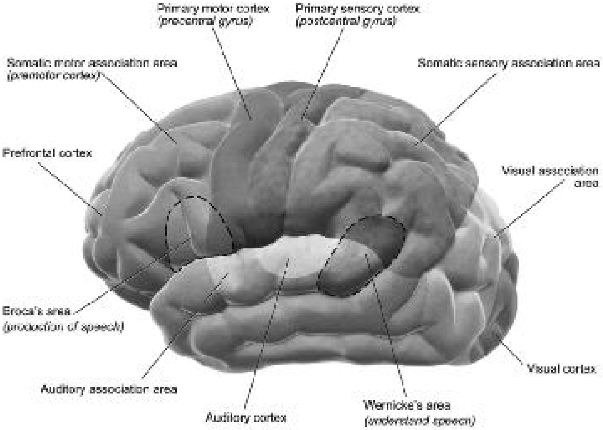

Primary motor cortex (precentral gyrus)
Primary sensory cortex (postcentral gyrus)
Somatic motor association area (premotor cortex)
Somatic sensory association area
Prefrontal cortex
Visual association area
Broca's area (production of speech)
Visual cortex
Auditory association area
Auditory cortex
Wernicke's area (understand speech)

Figure 16.2 Functional associations over the cerebrum

You can see from Figure 16.3 that the brain is divided into left and right cerebral **hemispheres**. These two sides don't mirror each other; rather, each typically specialises in different types of tasks (this is called **lateralisation**). Researchers do not yet agree on how specialised each hemisphere is, but the diagram gives you a general idea. Both Broca's area and Wernicke's area are generally located in the left hemisphere of the brain, though a significant number of people have the lateral functions reversed (both areas, along with other functions associated with logic and reasoning, are on the right), and there are others who do not show any such specialisation.

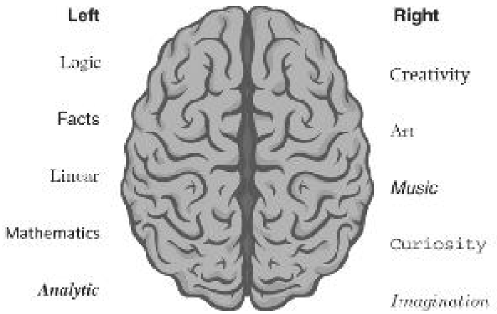


Figure 16.3 Lateralisation of the brain

One of the tests for uncovering which hemisphere controls language is something called **dichotic listening**. A person wears headphones and two different words are played into each ear, say the word *book* into the right ear, and *pen* into the left. What happens generally is that people hear more accurately what is played into the right ear because this is the ear that directly linked to the left hemisphere (the dominant side for speech). However, the reverse holds for non-linguistic sounds (let's say someone whistling or a toilet flushing). People will hear what is played in the left ear better than the right.

Only left-handers are in their right minds?

It used to be thought that left-handers had speech located in the right hemisphere – hence the old joke about left-handers being in their right minds (in possession of their faculties)! However, this is not the case: most lefties have speech controlled by the left hemisphere – in fact, around 90 per cent of the population do (estimates are usually between 80 and 95 per cent). So, some people have right hemisphere language specialisation and others have been found to have speech ability in both halves. It is true, though, that of those few people whose language is controlled by the right hemisphere, the majority are left-handed. The relationship between language and lateralisation is an extremely complex one, and remains the subject of continuing neurolinguistic research. There is a lot we don't understand, especially since much of what we know comes from clinical reports of people with brain damage or brain diseases. The relationship between handedness and brain specialisation is also complicated. As many of you will experience, even defining handedness is tricky – it isn't just a simple matter of which hand you prefer to write with.

The use of visual, rather than oral or aural signals, in sign languages is sometimes thought to mean that speakers of Auslan would process language on the right side of the brain alone. However, researchers have demonstrated that the impact of left-hemisphere damage in patients who communicate using sign languages is similar to that observed in spoken-language patients.

Interestingly, recent reanalysis of two of Broca's original samples using modern technology shows that there were also lesions in other areas of the brain in each of the two specimens. This indicates that difficulties with production in these two individuals may have been caused by problems in other areas of the brain. Furthermore, there are known cases in which Broca's area has been destroyed through illness or trauma, and some patients are still able to produce speech effectively. This shows that localisation in brain structures is not a simple matter of mapping one function directly to one location: the structure of the cerebrum and the way the brain functions are much more complex than that.

The cerebrum is made up of vast networks of billions upon billions of neurons that communicate with each other through electrical impulses; Figure 16.4 shows the structure of a neuron. Each neuron contains a cell body (also called the soma). It is surrounded by dendrites, which draw information into the neuron. Extending from the cell body is the axon, a series of cells covered in a fatty myelin sheath that connects to the synapses. The function of the synapses is to make connections to the dendrites of other neurons. The axons of some neurons are very long and may carry hundreds of synapses. The synapses send information to the neurons with which they are connected by either electrical or chemical means.

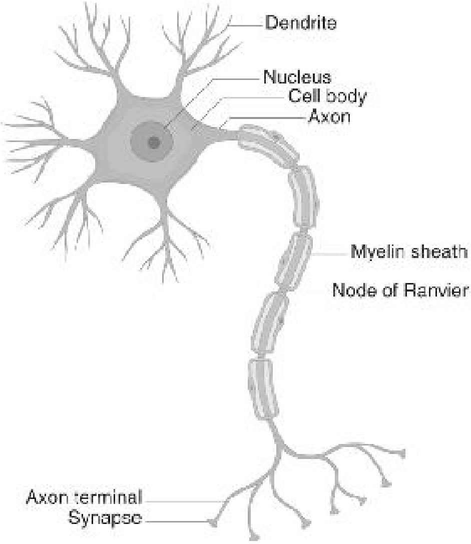


Figure 16.4 Structure of a neuron

The difficulties with assigning ‘language’ to any specific area of the brain are related to the multitudes of ways that neurons are able to organise and reorganise themselves. The system is more robust and flexible than it looks at first glance. The existence of people with brain injury in areas critical to language who are nevertheless able to communicate, is evidence that language functions are not simply or discretely localised. Probably the most famous case is that involving the construction foreman Phineas Gage. On 13 September 1848, an explosive was triggered prematurely during a detonation task and a tamping rod went straight through his skull. Miraculously, Gage survived, with his movement and speech intact. He did, however, experience considerable personality changes – as his friends reported, ‘Gage was no longer Gage’. His physician Dr Harlow described him as ‘fitful, irreverent, indulging at times in the grossest profanity (which was not previously his custom)’. The rod and Gage’s skull are now preserved in a museum at the Harvard Medical School, and with modern technology scientists have been able to reconstruct the path the rod took through Gage’s skull (Figure 16.5). This showed that most damage was inflicted on the frontal lobes, specifically something called Brodmann’s area – crucially, his lesion did not involve Broca’s area. However, the damage Gage sustained compromised his inhibition mechanisms and this explains his uncontrolled outbursts of swearing.

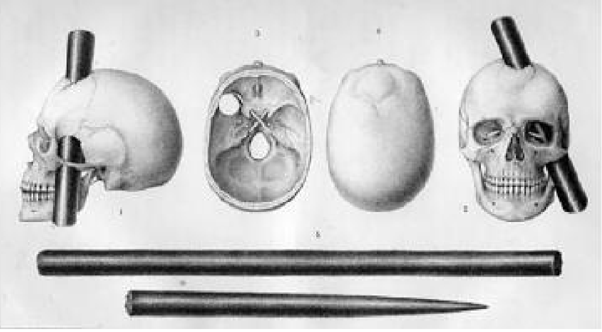


Figure 16.5 Phineas Gage's skull and tamping rod

Multiple ways to organise language – 'normal' need not be better

We have just observed that language tends to be more strongly associated with the left side of the brain, but that there are also individuals whose lateralisation associates language with the right side and others who do not display lateralisation to any notable degree. Some researchers have tended to make associations between right-side lateralisation and other non-typical types of behaviours (sometimes focusing on language alone, but often also on left- or right-handedness, as just described). For example, they have argued that lateralisation is an important stage in language development, speculating that it is associated with greater processing speed because information does not have to travel so far in the brain. Others have also claimed that lack of lateralisation to the 'correct' (that is, most frequently observed) side is an evolutionary disadvantage, reasoning from the idea that whatever is statistically normal is also biologically optimal.

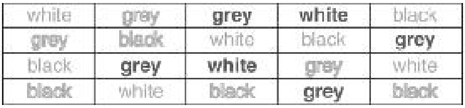
Yet, there is also plenty of evidence to suggest that atypical lateralisation doesn't result in abnormal linguistic proficiency. Researchers have measured performance on a range of language-related tests, including picture-verification tasks, word-verification tasks, engagement in artistic or academic pursuits and number of languages spoken, and there do not appear to be any strong correlations. In fact, one study suggests that the more unusual pattern of distributed language processing (bilateral language representation) actually has a significant benefit for an ageing population – an advantage that only becomes relevant to individuals later in life who experience something like a stroke. These people may be less likely to suffer language problems or may recover better from language impairment that occurred as a result of a brain trauma. As these researchers suggest, paradoxically, what appears to have started out as a disadvantage in the evolutionary past might now be a real advantage.

16.3 Ways of learning about language and the brain

So, if the brain is made up of billions upon billions of neurons that bind themselves into complex patterns of connection, while maintaining the superficial appearance of cold porridge, how have researchers figured out anything about the associations between structure and function? In this section, we provide an overview of how researchers have learnt about the relationship between language and the brain. The methods involved can broadly be divided into two types. First, there is a set of methods for learning about how language is organised within the brain by making inferences based on observable behaviour. These **behavioural methods** involve testing participants on a range of tasks with carefully controlled variables and thinking through the significance of the results. The methods in the second set, which have developed in more recent times, involve recording activity in the brain as indicated by changes in electrical activity or blood flow in different areas. These **neuroimaging techniques** allow researchers to explore the physical structure of the brain in fairly non-invasive ways. In both cases, experimental approaches normally are used – these approaches allow researchers to limit the amount of information with which they are dealing at one time by focusing on one specific question in each experiment.

Behavioural methods of psycholinguistic research

Most behavioural experiments in psycholinguistics rely on differences in the time it takes participants to perform particular tasks. The times involved are often miniscule, generally expressed in milliseconds. Harder tasks, which are understood to be tasks that involve more processing activity, take longer. For example, people often find it harder to read the word *blue* when it is written in red ink – something that researchers have established by timing readers' responses. People understand that the word says 'blue' more quickly if it is written in black ink. This is known as the Stroop effect (Figure 16.6). The participants' comprehension can be measured by timing how long it takes them to answer a question (which might simply be to read the word correctly out loud) or it can be measured by tracking how the participant's eyes move – how long they stay focused on the word. When participants have solved the task, their gaze moves away. The time responses take can also be measured against how accurate they are. Researchers can measure the speed–accuracy tradeoff between faster and slower responses to get an idea of how long a task takes to complete successfully. Although this is a great experimental method because it protects the findings from different types of bias, it requires very large numbers of responses across a large number of prompts. This is a significant cost that makes measuring speed–accuracy tradeoffs more expensive and time-consuming than other methods, meaning it is not terribly practical, given the resources that typically are available to researchers.



white	grey	grey	white	black
grey	black	white	black	grey
black	grey	white	grey	white
black	white	black	grey	black

Figure 16.6 The Stroop effect

The boggle effect

Taboo words are such attention grabbers. Look at the power of the designer label *FCUK* – even when we jumble up the letters, the word leaps out at us. Indeed, all sorts of psycholinguistic and neurological experiments verify that taboo words are more arousing, more shocking, more memorable and more evocative than all other language stimuli. A 2004 report by MacKay and colleagues describes the effects of emotion on memory and attention using the Stroop task. Taboo words were displayed in a salient colour, and participants were asked to name the colour and ignore the word. There were three significant effects.

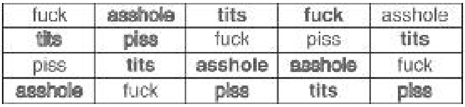
1 Participants were slower in naming the colour of the taboo words compared with the colour of the neutral words; we infer that they were distracted by the disturbing nature of taboo words (Steven Pinker describes this as 'an involuntary boggle'). Interestingly, this effect diminished with word repetition, which is consistent with what we have known for a long time – words wear out and emotional words are particularly prone to this.

2 Participants better recalled taboo words in surprise memory tests following colour naming. (Other experiments demonstrated that taboo words impair immediate recall of the preceding and succeeding words in rapidly presented lists.)

3 Participants more accurately remembered those colours that were consistently associated with taboo words rather than neutral words.

There is an anatomical dimension to all this. The processing of the emotional components of language, such as taboo words, belongs to what is known as the limbic system. This is a part of the midbrain (about the size of a walnut) and represents an older, deeper part of the mammalian brain. It is this limbic system that adds the emotional spice to the surrounding cerebral cortex (the part responsible for verbal reasoning, calculation, analytical thinking and rational thought).

Other behavioural techniques associated with psycholinguistic research include the following:



fuck	asshole	tits	fuck	asshole
tits	piss	fuck	piss	tits
piss	tits	asshole	asshole	fuck
asshole	fuck	piss	tits	piss

Figure 16.7 The boggle effect

- **Priming** is used to set up situations in which expected responses are more or less likely. For example, it is easier to read the word *soup* if the word *broth* is presented first. The word *chicken* has a weaker but still positive effect for the word *soup*.
- **Gating techniques** measure how much acoustic information is required before the participant can correctly identify a word. For example, how much of the word needs to be available before a listener can distinguish between *part* and *park*?
- **Eye tracking** shows that there are systematic patterns in the way readers' eyes track across written text. Reading involves a mixture of fixations on particular letters and saccades (or scans of neighbouring text). In fluent readers, nine out of ten saccades involve scanning forward to new text.

Chilling the blood and raising gooseflesh

> There is something very terrible in an oath torn from its proper home and suddenly implanted in the wrong social atmosphere. In these circumstances the alien form is endowed by the hearers with mysterious and uncanny meanings; it chills the blood and raises gooseflesh.
>
> Henry Cecil Wyld (1920), *A History of Modern Colloquial English*

In several places in this book, we have described words as symbolic (the correlation between their form and meaning is arbitrary). Exceptions to this are sound-symbolic words (*Woof! Whoosh!*). Another group of exceptions comprises the so-called dirty words. Think about it: if words are arbitrary sounds that we hear, or letters that we see on the page, how can they be 'dirty'? Taboo words are somehow felt to be intrinsically nasty – it is as if the form of their expression communicates the essential nature of what it is that they represent.

This is shown nicely by a group of experiments using a technique known as electrodermal monitoring: participants are hooked up to fingertip electrodes that then measure their emotional responses to a range of taboo words. It is a kind of polygraph testing that records what in the trade are called 'galvanic skin responses' or 'skin conductance responses' – and 'dirty' words evoke stronger responses than other words. These might sound much like 'goosebumps', but in fact they aren't. Galvanic skin responses (GSR) are changes in the skin's ability to conduct electricity and are triggered by changes in emotional stimulus (pleasant or unpleasant) – they might involve fright, anger, shock or sexual feelings. Goosebumps (GB) refer to the visible roughness of the skin (resembling that of a plucked goose), which can develop in response to cold, but also strong emotions such as fear or sudden excitement. So, on the one hand, we are looking at an increase in sweat gland activity (the GSR) and, on the other, hairs standing on end (the GB). They are two quite different beasts, although the feeling that most people think of as 'creeping of the flesh' can, in a non-specialised sense, apply broadly to both states – hence Wyld's description in the above quotation where he refers to the ability of swear words to chill the blood and raise gooseflesh. Wyld's account (while a little over the top!) now has plenty of scientific confirmation.

Neuroimaging techniques

Neuroimaging techniques allow us to ‘see’ something of what is happening inside the brain as it processes language. These techniques represent an enormous progression from the early days of neurophysiology, when autopsy was the main method available for understanding brain structure. The following three methods of neuroimaging have been used to further our understanding of how the brain processes language.

- **Electro-encephalography (EEG)** records levels of electrical activity through the scalp. The electrical activity is associated with bundles of neurons within the brain firing (sending impulses to each other).
- **Functional magnetic resonance imaging (fMRI)** measures blood flow in the brain. This technique is very useful for showing which parts of the brain are activated during different tasks.
- **Magneto-encephalography (MEG)** uses magnets to measure electrical activity within the brain. This technique is relatively recent and requires more expensive machinery than EEG, but it has the advantage of providing useful information about both the timing and the location of brain activity.

Colours, brain hemispheres and language

Following on from research like the ‘Russian blues experiment’ reported above, researchers have sought to explore whether there are hemispheric effects associated with colour-identification tasks. Tan and colleagues constructed a task that required participants to decide whether a pair of colour squares flashed in front of them contained a pair of matching colours or different colours. Participants completed this task while their brain activity was being scanned using MRI technology. This meant that the researchers could locate and measure activity in the brain and see which areas were involved in the task.

The researchers found that completing the task with easy-to-name colours (red, green and blue) was associated with increased activity in two specific areas of the left hemisphere. When the task involved more difficult-to-name colours (samples that might be labelled burgundy, teal and beige), these areas of the brain were not so heavily activated. The researchers also demonstrated that the areas activated strongly by the easy-to-name colour-matching task were the same areas activated when the participants were asked to name or read the names of the colours. In other words, even if language is not required for a colour-naming task (because the task is just to identify colours as either similar or different), the parts of the brain associated with naming colours in the left hemisphere will be triggered for easy-to-name colours.

16.4 Models of language processing

On a physical level, then, we can know things about how neurons communicate with each other through electrical and chemical impulses, and we can observe which parts of the brain are active, and for how long, in relation to particular tasks. This is still a fair way from providing us with detailed knowledge of language production or comprehension. Given the complexity, adaptability and variation within the human population in terms of brain functioning, though, there may not be much in the way of a single truth to be discovered. Models of language production and comprehension are a useful response to all of this complexity. They provide us with a way forward in terms of articulating what we currently understand in more general terms and we can apply these models to real-world language issues.

There are two basic approaches to modelling language processing: **top–down** and **bottom–up**. In these models, the 'higher' spaces are associated with meaning-making, while the 'lower' spaces are associated with physical representations of language, such as sounds. These two processes are described in Figure 16.8.

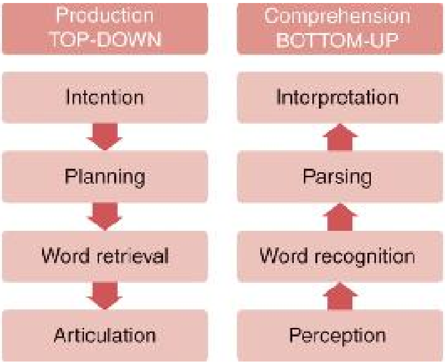


Figure 16.8 Top-down and bottom-up models of language processing

It probably won't surprise you to learn that things are not as simple as these diagrams suggest. Although there is some broad truth in these models, in practice, the processes involved in both production and comprehension are more **interactive**. This is clear, for example, in the gating techniques mentioned above where the hearer attends to the signal (the sound of the word) and also makes predictions about which word it will turn out to be. In this example, perception is happening at the same time as attempts at word recognition are also being made.

The challenges of understanding spoken language

Figure 16.8 gives you an inkling of the very complex cognitive processes involved in understanding spoken language, particularly when we recall that this process is interactive. This means that the brain is juggling a number of tasks simultaneously and at great speed. Children who have difficulties with understanding spoken language provide us with insight into the complexity of this task. Speech pathologists and teachers who work with these children have identified a number of areas where intervention can help.

First, here are some symptoms of difficulties with understanding spoken language:

- difficulty answering questions
- difficulty following instructions
- difficulty remembering what has been said
- difficulty maintaining attention.

Children who are struggling to comprehend speech work harder than their peers – yet, their overall performance on tasks that depend on language comprehension remains poorer than that of their peers. Excluding hearing loss problems, children can struggle with comprehending speech for a whole range of reasons, relating to:

- ability to focus on speech sounds as a sub-set of the ambient noise in the area (problems with perception)
- ability to relate sounds to words and identify words correctly (problems with lexical retrieval)
- short-term memory problems (holding information long

enough to get through the processing tasks involved in comprehension).

Interventions can include actions that support the child's ability to hear more clearly. For example, seating children at the front of the classroom and providing slower speech from adult caregivers can help overcome some basic difficulties that make perception more challenging. Activities that address particular processing issues are also useful. For example, phonological awareness training involves practice in manipulating sounds in words (What's a word that rhymes with *cat*? How many sounds in *house*?). This supports the development of word-recognition skills by providing the child with practice distinguishing between different words based on their sound profiles.

We will explore language production and comprehension processes in the following sections, beginning with language production and returning to language development issues later in the chapter.

16.5 Language production

Language **production** involves planning at a range of levels as speakers assemble their thoughts and prepare to express them verbally. The following stages are often identified in language production:

1 conceptualisation – speakers form an idea of what they want to say

2 grammatical and lexical planning – speakers (unconsciously) organise words into an appropriate grammatical framework

3 phonological planning – speakers identify the form of the utterance

4 phonetic adjustment – speakers may simplify the utterance to make articulation easier

5 articulation – speakers utter their message.

Error analysis

Error analysis is one fruitful and non-invasive way that researchers can learn about language production processes and issues. Pauses and **hesitations** have been one focus of research, while transpositions of words, morphemes and sounds have been another area of interest.

Pausing can be used deliberately for rhetorical effect, or to signal places in an interaction when speakers and listeners can organise their turns. Politicians often use fairly long pauses to make a point and provide time for listeners' reactions (unless they are being interviewed on television or radio – where silence is dead time). However, pauses that are associated with the speaker struggling to organise their speech production clearly fall short of an ideal delivery. These pauses and hesitations provide a range of useful bits of information about how the planning process happens. This is because they reflect places where speakers have to make choices about how they intend to communicate their message. Furthermore, longer delays tend to correspond to larger or more complex planning problems. Hesitations – that is, vocalisations such as *um, er* and *oh* – signal to listeners that speakers know their delivery is falling short of the ideal and that they are working hard to rectify this. In other words, not only are speakers planning and performing their speech, they are also monitoring their performance and making adjustments to it as they go along. This is further evidence that processing happens in an interactive way during an utterance. Mispronunciation, for example, may trigger correction, including a loop back through intention, planning, word retrieval and articulation. Word retrieval is also a significant trigger for hesitation.

Speech errors also provide information about planning. Many errors involve the reordering of sounds and words. These types of changes are only possible if the utterance has been planned before it is spoken, providing researchers with evidence about the timing and organisation of speech. There are two ways of collecting speech error data. One is noting errors that occur in the world around us. The other is to induce errors in participants in laboratory experiments. Speech errors can occur at any level of linguistic structure. A series of examples based on the target sentence *She planned to dig up some potatoes for her tea* are provided in the following list. We have bolded the words involved in each case.

- Phrase (exchange): *She planned to dig up* ***her tea*** *for* ***some potatoes***.
- Word (substitution): *She* ***phoned*** *to dig up some potatoes for her tea.*
- Word (exchange): *She planned to dig up some* ***tea*** *for her* ***potatoes***.
- Inflectional morpheme (shift): *She planned to dig up some*

potato for her ***teas****.*

- Morpheme (exchange): *She* ***dug*** *to* ***plan*** *up some potatoes for her tea.*
- Syllable onset (preservation): *She* ***planned*** *to dig up some* ***plotatoes*** *for her tea.*
- Phonological feature (anticipation or preservation): *She planned to* ***dig*** *up some potatoes for her* ***dea****.*

An interesting observation about word exchange is that it typically involves words from the same grammatical category (for instance, the nouns *potatoes* and *tea* in the example above). This pattern of error suggests that there is a moment between identifying the appropriate sentence structure and words, and then 'slotting' the words into the right places where error can arise.

Spoonerisms

One of the more entertaining categories of error is known as the spoonerism. The name was inspired by the Reverend William Archibald Spooner, who worked at Oxford in the late 1800s (he became well known for his (real or alleged) slips of the tongue). As described in Chapter 8, these spoonerisms involve the exchange of the initial sounds of two words – interestingly, they never involve nonsense words. They are sometimes also constructed deliberately because the results can be quite funny (and we suspect some of Reverend Spooner's slips were the creation of his imaginative students). The following examples are attributed to Spooner:

- *Three cheers for our queer old dean!* (… dear old queen, referring to Queen Victoria)
- *Is the bean dizzy?* (Is the Dean busy?)
- *You have hissed all my mystery lectures. You have tasted a whole worm. Please leave Oxford on the next town drain*. (You have missed all my history lectures. You have wasted a whole term. Please leave Oxford on the next down train).

A selection of other popular examples is provided below.

Intended utterance	**Actual utterance**
go and take a shower	go and shake a tower

blow your nose	know your blows
pack of lies	lack of pies
picking your nose	nicking your pose
ease my tears	tease my ears
you have very bad manners	you have very mad banners

Spoonerisms can even be induced in the laboratory using the sorts of priming techniques we mentioned earlier. Participants are presented with word pairs that they read silently, but occasionally are required to say. Spoonerisms are sometimes induced by the phonological makeup of preceding pairs of words, and researchers believe these errors occur because two competing speech plans are activated; and speakers aren’t able to block the flawed plan prior to articulation.

Lexical retrieval

Lexical retrieval has been a particularly active area of study in psycholinguistics. The goal for speakers is to retrieve and articulate the correct phonological form of the word by matching a form to their intended meaning. The Spoonerism examples show that sometimes the start of the word can get jumbled up with a neighbouring word to humorous effect. But, there are other issues for researchers to consider. Speakers of inflectional languages, for example, need to be able to sort out which form of the word is appropriate to the context. You can see the strength of these aspects of word knowledge when you consider the following two sentences (repeated from above).

(16.1) Target sentence: *She planned to dig up some potatoes for her tea.*

(16.2) Stem morpheme (exchange): *She dug to plan up some potatoes for her tea.*

The target sentence has a past tense verb with the past tense suffix *-ed* and an infinitive verb in the subordinate clause. When the two verbs PLAN and DIG are reversed, the past tense form for *dig* is nevertheless correctly inflected as *dug*. When the lexeme DIG is retrieved, it brings along information about how it should be used in a past tense context.

Tip of the tongue (TOT)

One widely experienced aspect of speech production is the tip of the tongue (TOT) phenomenon. In this state, the speaker has trouble recalling a word even though they have a very strong sense of its shape and meaning. You must have experienced this at some time: you can't actually supply the word but you can give semantic information and often a lot of phonetic detail. Experiments have even been devised to induce a TOT phenomenon. Participants might be asked to supply words corresponding to particularly tricky definitions and, when they can't supply the actual word, they are asked to give as much of the information as they can about it. It turns out people are usually pretty good at recalling the beginnings of the words (about 70 per cent accuracy), as well as the number of syllables and the position of the stress. Typically, their recall of word endings is better than their recall of middles (possibly because suffixes in English carry a lot of important grammatical information). So, beginnings and endings are imagined more clearly that the middles, and psycholinguists like Jean Aitchison have dubbed this 'the bathtub effect' (true, it's not the usual way we sit in baths – you have to think of a cartoon character in the bath – see Figure 16.9).

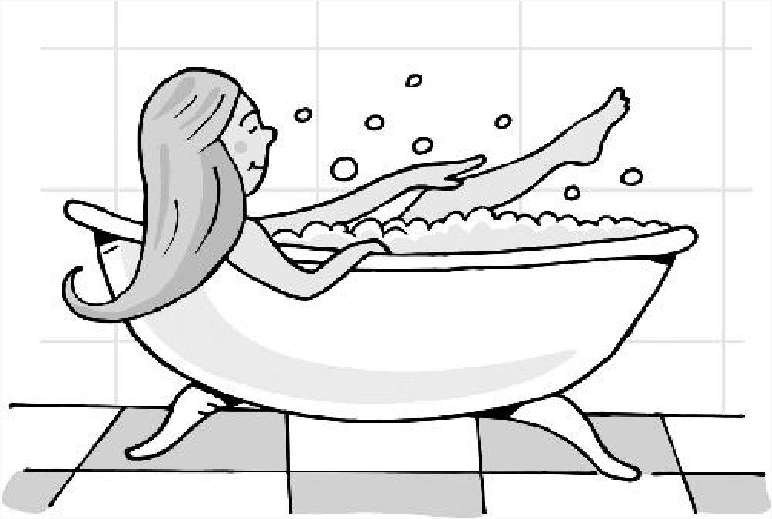


Figure 16.9 The bathtub effect

16.6 Comprehending words

> The goodness of the true pun is in the direct ratio of its intolerability.
>
> Edgar Allen Poe, 1844–49, *Marginalia*

Research on language comprehension has tended to split off the task of identifying words from the task of interpreting sentences, even though it is also acknowledged that these processes are interactive. We will focus here on words and consider sentence comprehension below.

The first challenge we face as listeners trying to comprehend someone else's speech is the need to convert the sound waves they produce into linguistically significant sequences of phones (sounds) or phonemes (see also Chapters 7 and 8). One strategy that appears to be used widely across many languages is the **Possible Word Constraint**. This constraint means that the listener will work to assign every segment of speech to a word. Leftovers are to be avoided. In a sequence like [ðæɪmetəfoːθtɑem], this prompts listeners to avoid identifying some possible words and preferring others that avoid phonological residue. The sequence contains all of the following potential words: *they, aim, met, metaphor, a, for, fourth, time* and *I'm*. The only set of words that uses all of the phonological segments in the signal is *They met a fourth time*. This is the segmentation that listeners readily prefer.

The Possible Word Constraint is useful, but it doesn't solve every problem the listener might face. Recall the discussion of mondegreens in Chapters 6 and 7. These are examples where listeners have identified alternative but complete strategies for assigning the phonological segments into words, perhaps with occasional reassignment of one or two phones. For example, *The ants are my friends* makes pretty thorough use of the phonological segments available in Bob Dylan's *The answer my friends*. The oronyms we discussed in Chapters 7 and 8 are similar – for example, *We're a right pair* versus *We're a ripe pear*. This misinterpretation is facilitated by the non-release of the final voiceless stop in *right* [weːɹəɹɑeʔpeː]. These examples become feasible to the listener when there is nothing in the context to force a correction. It makes sense that song words and brief remarks are associated with this type of problem since brief remarks have minimal linguistic context (that is, they are not part of larger structures and rely on shared inferences for their interpretation), while song lyrics are poetic and lead the listener to accept sometimes surprising word combinations. (*I'll never be your pizza burning* doesn't seem so odd from The Rolling Stones, though the original was actually *I'll never be your beast of burden*).

Suggestibility and the risks of wrongful convictions

Our experience with mondegreens and mishearings shows that indistinct audio is open to all sorts of multiple hearings – and the more indistinct, the more open the audio is to alternative interpretations. In Chapter 7, we emphasised the important role listeners' unconscious expectations play in shaping speech perception – prior exposure to a word or to contextual cues can then prod hearers to confidently hear words that were never actually uttered. The work of forensic linguist Helen Fraser plainly shows the potentially catastrophic fallout of these priming effects. As she describes it:

> The way a sound can go from meaningless static to clear phrase under the influence of a prime is exactly what happens with lots of indistinct covert recordings [conversations lawfully captured by police without the knowledge of the speakers], like the one that got a man convicted for a murder based on words he didn't say.

Research going back even as far as the 1920s shows that listeners are highly susceptible to being primed by suggestions about what they might hear. And yet, transcripts of covert police recordings continue to be presented to a court without any acknowledgement of this research.

Clearly, juries need a transcript to help them interpret poor-quality covert recordings, but there is much that needs refining in relation to the current practice of using police transcripts to aid the interpretation of indistinct audio. The only people who can properly assess these transcripts before they are used in court are those with independence and expertise in the relevant branches of phonetic, forensic and cognitive science. As Fraser emphasises, leaving it to the jury is not fair on jurors – and it is certainly not fair on defendants.

Other cognitive processes that support correct word identification include frequency (as described in relation to gating techniques in section 16.3) and **neighbourhood effects**, which refer to the distribution of words that are similar to the target word, differing by only one segment. A target word like *bin* has lots of neighbours (words with the same initial consonant and vowel) – *bit, bill, bing, bid, big*, and so on. Words that exist in dense neighbourhoods, like *bin*, take longer for listeners to identify than words (like *keg*) with few competing forms. Listeners also have to be able to deal with **lexical ambiguity** – the existence of homonyms (such as *grouse* 'fantastic' versus *grouse* 'type of game bird'), and polysemous words (such as *glass* = 'beaker', 'mirror', 'barometer') that have multiple possible senses and must be worked out with reference to the sentence context. The grammatical features of each word also need to be determined (for any given utterance, is *love* a noun or is it a verb?) as part of the process of sorting out how they relate to the overall message. Listeners, like speakers, employ monitoring techniques to ensure they are processing language effectively. Much research on sentence comprehension has been concerned with the strategies that listeners use to resolve potentially ambiguous sentences. It has often involved participants reading from cues and providing evidence of their reading speed at different points in the reading process.

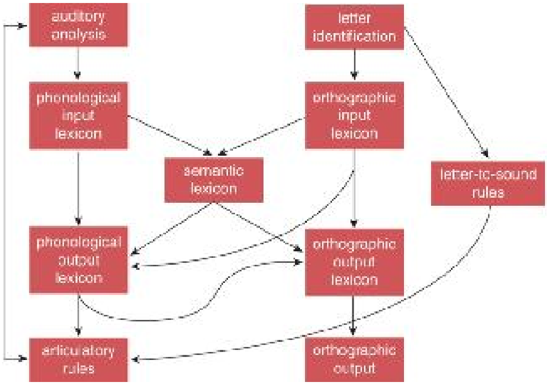


Figure 16.10 Possible model for the comprehension of both spoken and written words

Mind your p's and q's

Reading is a relatively late development in human evolution, and it is fascinating to consider how we have been able to co-opt brain functions originally used for other visual tasks to support reading processes. One of the most amazing things about this transition is that we have had to overcome our natural ability to recognise objects from any orientation (we know an elephant from either end) since in writing the same shape in a different orientation is actually a different letter (consider *b, d, p* and *q; n* and *u*). Learning to read requires us to learn to convert visual information into imagined auditory information. Our internal representation of a word is at first mediated by our knowledge of the phonology represented by the letters in the word. As we become more skilled readers, the process of word recognition requires less phonological analysis, making word retrieval quicker and easier.

16.7 Comprehending sentences

Each time a candidate word is identified, the next task for the reader or listener is to figure out what relationship the word has to the rest of the message. **Garden-path sentences** are used to explore the ways in which readers assign meaning to phonological units they have identified as words. These sentences are constructed to trigger an early and plausible, but incorrect, interpretation from the reader (here you have to imagine someone being misled and lured down a path that is a dead end!). The following two examples show garden-path sentences. (In spoken language, prosody can force a particular interpretation and, thus, disambiguate the meaning.)

(16.3) I persuaded her children are adorable

(16.4) Because I read the second hour of waiting went quickly

The sentence in example 16.3 is difficult to parse since the sequence *her children* is first interpreted as a single unit (a noun phrase). This analysis has to be revised in order to complete the sentence correctly once the second verb *are* and the adjective *adorable* are taken into account (at the completion of reading the sentence). The subject of the second clause in example 16.3 is *children*, while the object of the first clause is the pronoun *her*. Similarly, in example 16.4, the listener is likely to assume that *the second hour of waiting* is a book title – as anticipated by the word *read* – and will only seek an alternative interpretation once the final phrase *went quickly* becomes available. In this case, the entire noun phrase *the second hour of waiting* is reassigned as the subject of the second clause. Try reading these sentences out loud to friend or family member and see how they do a double take. Experiencing garden path sentences can be even more confusing when listening rather than reading, because of the processing constraints imposed by having to interpret in real time.

What's going on here? It seems that as people parse language (spoken or written), they assign each word they encounter to a particular slot in the structure of a clause. Furthermore, they only begin to construct new clauses when it is clear that the current clause is 'full' or some other information, such as an unexpected 'extra' verb, appears (as in the examples above). The reanalysis places significant demands on short-term memory because the listener or reader has to go back and undo some of their established understanding, while continuing to parse the new information that is coming in.

Assigning a sensible structure to the string of words coming in does not represent the final stage in the process of language comprehension. There is one remaining issue for the listener or reader to resolve: what was the intention behind the message? For example, we noted in Chapter 14 that a single sentence type can express all sorts of meanings. Speakers use declarative sentences as statements (*I'm cold*), questions (*You look cold*?) or even (subtle) commands (*It's cold*, which might be a polite request to shut the window). Sarcasm and irony have a significant impact on how the reader or listener should finally respond to the message.

16.8 The bilingual brain

We turn now to consider a range of ways in which the language profiles of individuals can vary, beginning with research on how individuals organise multiple languages in relation to each other. As discussed in Chapter 12, bilingualism refers to a person's ability to communicate effectively in two languages. In some publications, it is also used as a cover term for multilingualism – speaking three or more languages. Although there is a widespread belief (especially in English-speaking communities) that monolingualism is more normal and less problematic than bilingualism (or multilingualism), for the majority of the human population, speaking more than one language is a normal ability. Moreover, research over the past 50 years has established that there are many cognitive advantages to the ability to speak more than one language. This includes evidence that bilinguals may have greater **metalinguistic awareness** (the ability to notice things about language) and enhanced skills in being aware of how they are thinking (**metacognitive skills**). Bilinguals have also shown stronger symbolic representation and abstract reasoning skills, better learning strategies, stronger problem-solving skills, greater success with theory-of-mind tasks (see below), enhanced creative and divergent thinking skills and greater cognitive flexibility.

In other words, many recent studies show that bilinguals are better at thinking creatively than monolinguals. This is thought to be because acquiring more than one language fosters increased fluency, flexibility, originality and elaboration in thinking – insofar as these can be measured in psychological tests. Although the reasons for this difference have not been clearly identified, it seems that two languages provide the speaker with additional conceptual resources. Speaking more than one language gives people experience in viewing the world from more than one perspective. This mental strategy can be used in creative thinking and problem-solving as well as speaking. The brain scans of bilingual individuals are interesting in this regard; new research into neurobiology of bilingualism has been showing up greater grey-matter density (brain tissue packed with information-processing neurons and fibres) in the brain's language-dominant left hemisphere. The density appears to be most obvious in those who are very proficient in a second language and those who have learned a second language at a young age (before the age of five years). Studies also show that the bilingual experience might even create what has been called a 'cognitive reserve' that helps to delay late-onset dementia.

So, why is there such a gap between public assumptions about the capacities of bilinguals and the research? One reason is that current research follows on from an earlier series of studies that found that bilingualism disadvantaged speakers in a range of different ways, including giving them smaller vocabularies. This finding reflected the fact that vocabulary was only counted for one of the languages spoken – generally the dominant language of the wider community. Knowledge about minority language(s) was not included in the test. Since most bilinguals have somewhat different sets of vocabulary in each language (a reflection of the different communication tasks associated with them), these tests were not measuring each individual's total linguistic ability. A second series of tests reported that bilingualism stunted people's cognitive abilities, reflecting the fact that bilinguals performed worse than their monolingual peers on IQ tests. The IQ tests were often administered in the dominant but less preferred language of the test participants, meaning they did not provide participants with the opportunity to really show what they could do. There were other serious methodological problems with these studies. For one thing, people were often identified as bilingual not on the basis of their ability to speak two or more languages fluently, but by applying criteria such as their residential postcode, the country of birth of their parents or even the form of their surname!

It was not until the 1960s that research on bilinguals began to take issues of methodology seriously. After controlling for age, gender and socioeconomic status, and applying consistent measures of bilingualism, testing methods improved significantly and the results changed dramatically. Since that time, research has shown increasingly that bilingualism is in fact an advantage both in terms of cognitive development and cognitive resilience. Although there is great potential for cognitive advantages associated with bilingualism, not all children who are raised bilingually are able to benefit, given the broader social context. To take advantage of the cognitive benefits of bilingualism, children need to be able to learn and develop in environments where their identity as speakers of other languages is recognised and valued; socioeconomic status (a confounding factor in early studies) continues to be a significant issue for many bilingual speakers of minority languages today.

Where we store taboo words: Bilingual evidence

Earlier we presented you with the experimental evidence showing that taboo words are processed differently from ordinary language – they have more accurate recognition and recall. A slightly different perspective is provided by experiments that examine the emotional responsiveness of bilingual speakers.

For example, a study reported by Ayçiçegi-Dinn and Caldwell-Harris of 32 English–Turkish bilinguals used fingertip electrodes to measure the emotional responses of participants to taboo and other kinds of highly charged words. Researchers found that the greatest emotional reactivity was to taboo words in both languages. Consistent with earlier studies using only monolingual speakers, taboo words (for example, curses and sexual terms) generated far stronger responses measured by amplitude of skin conductance than other kinds of emotional words, such as reprimands (*Don't do that!*), aversive words (*cancer, kill, death*), positive words (*bride, joy, kind*) and neutral words; in fact, the amplitude was almost double that of neutral words. Even stronger responses were reported for the speaker's first language. A number of participants during the debriefing sessions reported they 'felt nothing' when they heard or even uttered the taboo expressions in their second language.

So, why would the language learnt early in life – especially emotive expressions like 'dirty' words – elicit stronger physiological responses than language learnt later in life? As some of you have probably experienced, children pick up the emotional meaning of these words very early on, well before they have knowledge of the social and cultural pressures that suppress them. Taboo words come with social rewards (children soon learn that they get maximum attention if they use them), and they come with penalties (reprimands, even banishments: 'Go to your room!'). So, when we learn taboo words at a young, impressionable age, they are presumably coded with the equivalent of a linguistic health warning: dangerous – use only with extreme caution. As our brain systems mature and we develop discretion about what we say, we then learn to bury taboo words and phrases via the usual neurological processes of inhibition. However, expressions acquired later in life will lack the childhood aura of the 'forbidden'. Without these emotional associations, they have quite different representations in our brains. Not surprisingly, then, 'dirty' words in a language learnt later in life trigger weaker emotional responses and won't cause the flesh to creep in the same way. They don't have such a special place in the bilingual brain.

16.9 Language disability

Language disability has always been a significant focus of research on language and the brain. As we noted in section 16.2, language disorders provided early researchers with evidence for specialised areas of the brain relating to language. In this section, we review some of the more common types of language disorders that are experienced and refer to the typical treatment or ameliorative responses currently available. Although voice disorders are a focus of speech and language pathologists (for example, stuttering), we will not explore them in detail here – we focus particularly on problems related to language as a system of communication more broadly. Before turning to consider a range of language difficulties in more detail, it is useful to make a couple of preliminary distinctions. For this reason, we have organised this material into two sections – focusing first on language disabilities that emerge as the child develops and then turning to language disability that emerges later in life as the result of damage to the brain.

As you read this material, it is worth bearing in mind that speech and language pathologists have become increasingly aware of the potential difficulties that labelling patients with a specific diagnosis can pose. Benefits accrue when an accurate diagnosis results in access to appropriate and effective treatment; however, negative results can be manifold. Incorrect labelling has sometimes led to children who are developing normally but are not able to perform in the same ways as their peers at school (perhaps for cultural or socioeconomic reasons) having reduced access to the regular curriculum and the opportunities it presents. Steps that would allow these children to adjust to school are not taken – the problems they experience become theirs rather than failures of the school system to meet their needs. Incorrect diagnosis is also an issue when it leads to problems being missed. This results in individuals with real needs for language-specific interventions receiving inappropriate treatment – sometimes for many years. But labelling has another cost, and that is to place the burden of the problem on the individual who is so labelled. When treatments and interventions fail or when the problem seems to get worse, it is easy to blame the person with the diagnosis for the problem rather than persisting with attempts to address the issues that led to diagnosis in the first place. Especially in the context of developmental disability, it is also important to acknowledge the impact that these labels have on the developing child's self-image and long-term prospects.

Tourette's Syndrome

Earlier, we described some of the evidence linking taboo language to the limbic system (the emotional processing areas of the brain). Further support for this association comes from people manifesting Gilles de la Tourette's Syndrome (TS), a neurological disorder associated with frequent and repetitive involuntary vocal tics (noises or utterances) and motor tics (muscle movements). The range of symptoms that can be seen in TS includes *coprolalia* (the compulsion to vocalise obscene or other socially unacceptable words or phrases). The tics are a type of spasm that can be more or less controlled (much like an urge to scratch an itch or to sneeze). The tics increase with stress and decrease when the person is relaxed or concentrating on an absorbing task; some people can hold them back for hours at a time, although this sometimes leads to more severe outbursts later.

A number of researchers have characterised TS as a breakdown of the brain's inhibitory mechanisms; in other words, the characteristic tics are understood to occur when there is a failure to inhibit bits of 'normal' behaviour. There is a substantial body of data based on post-mortem and neuroimaging studies that implicates the basal ganglia (the interconnected areas of grey matter located deep below the cerebral cortex) both in TS and related neurological disorders. These basal ganglia are the hidden pathways that contribute to the multiple circuits involved in the different cognitive and limbic processes that regulate emotions, thoughts and movement; neurologists have associated TS with a failure to suppress sub-sets of these mini-circuits. For example, the outbursts of foul language would be associated with a failure to inhibit portions of the limbic mini-circuits.

It seems that the obscenities associated with this behavioural disorder involve whatever is most socially inappropriate for the time. There is one case of a nine-year-old boy who lost his coprolalia entirely when he was told the words he was uttering were not obscene. A bilingual English–Spanish speaker reportedly switched the language of his outbursts depending on the language being spoken around him. So, while taboo language is rooted deeply in human neural anatomy, it is the sociocultural setting that clothes the emotion with expression – or, to quote the mother of a five-year-old girl with TS, ‘society shapes the noise that is made’.

Developmental disability

It has become increasingly clear that genetics may have a significant role to play in language-related developmental disabilities. Many children with difficulties in a range of language development areas (late onset of speech, reading problems, difficulties with understanding caregivers) have relatives who experience similar problems. Furthermore, it has also become clear that most language disabilities that emerge during development are associated with either cognitive or hearing problems, or both. In fact, it can be difficult for researchers to tease apart the causes and effects in this area since cognitive problems have such pervasive effects on development, and hearing problems impact on communication in terms of **reception**, learning and production.

Developmental disabilities refer here to disabilities that become apparent as the child develops at a significantly slower rate than their peers. Child-raising books rightly emphasise that there is a significant range for normally developing children. For example, we can expect first words any time between six and 12 months. Alternatively, some twelve-month-olds can say one word, while others have a vocabulary of 50. Both these figure, and every figure in between, is within the range expected in a healthy developing child. When children show significant delays in their language development (and these may be associated with other areas of concern, relating to their ability to interact socially, for example), they may eventually be identified as having one of the following language disabilities. (Other language difficulties not described below but identified with language development include pragmatic impairment, learning disabilities and reading impairment.)

- **Autism spectrum disorder (ASD):** this may include difficulties in social interaction, communication and aversive behaviours (that is, behaviours as diverse as rocking the body, squeaking or insisting on certain routines). Communication difficulties for individuals with ASD vary widely – with the exception of children with Asperger's disorder, language typically is delayed. All individuals with ASD display difficulty with interactional and pragmatic aspects of interpersonal communication, such as greetings and leave-takings, effective conversational skills and the ability to take the views of others into account. This may have a significant impact on social interaction.

- **Delayed language development:** this manifests largely in the preschool years (two to four years of age), and involves difficulties with understanding, using and producing language. Some children catch up with their peers by the time they start school (after the age of five), while others do not. Those children who continue to show difficulties with language development may be identified as having a specific language impairment (see below). The key diagnostic feature of children who are included in this category is that there is no other associated problem that can account for the language delay.
- **Specific language impairment (SLI):** this refers to children with language delay who have not caught up with their peers by primary-school age and do not have other problems that can explain the presence of the language delay. Researchers have been unable to identify straightforward causes for SLI; this is a task made all the more difficult because the label SLI covers a wide range of different disorders. It is therefore not surprising that interventions for children in this group vary considerably.

Communicating imagined worlds

One area that has been identified as challenging for children with autism is known as **theory of mind**. This is the capacity of individuals to recognise and respond to the fact that other people have minds and thoughts that are different from their own. A noteworthy feature of play in young children is their enjoyment of pretend play. (For example, a potential adult playmate might be told: 'You be the baby and I'll be the mummy and we're going shopping … '). Successful pretend play requires that all participants have a shared understanding of the scenario for the play and how it should proceed. To get this right, playmates have to talk about how the game will happen.

Researchers distinguish between talk about the game that is ‘out of frame’ (when participants stop playing, resume their normal identities and talk about the game) from ‘within frame’ talk (when participants try to maintain their play roles while negotiating: ‘Now Baby, you get in the car and I will drive you to the shops.’). Children often prefer to negotiate about the game within their roles (that is, within frame).

Children with autism have sometimes been described as being unable to participate in pretend play, although recent studies have shown that this is not necessarily the case. Australian researchers Susan Douglas and Lesley Stirling explored the ways in which children with autism do in fact engage in make-believe play, focusing on the types of negotiation between partners that could be observed. They noticed that children with autism find it more difficult to shift out of frame than typically developing children, but more importantly, it is challenging for them to participate in jointly developing a shared pretend frame that everyone plays within.

Acquired disability

There are various pathways to language disability aside from developmental issues. One is when language disability results from trauma to the brain resulting from stroke or brain injury. Other types are associated with more gradual loss of brain function through the development of tumours or degenerative processes like dementia. These are characterised by progressive or chronic impairment of a number of cognitive functions, including memory, critical thinking and orientation, and deteriorating emotional control and social behaviour. Dementia in particular is associated with a number of different medical conditions, such as Alzheimer’s disease, Parkinson’s disease and Huntington’s disease.

In Chapter 5, we briefly introduced you to **aphasia** (from the Greek ‘without speech’) – the difficulties in producing fluent, meaningful speech associated with injury to the left hemisphere of the brain (in patients with typical lateralisation).

Aphasia covers a range of communicative problems, including the following.

- **Anomia** refers to difficulties with finding the right word.
- **Apraxia** refers to speech that is non-fluent – slow and halting,

and lacking typical prosody.

- **Semantic paraphrasia** involves the selection of the wrong word – for example, *bus* for *car*.
- **Phonemic paraphrasia** involves the selection of an incorrect phoneme, such as *foot* for *put*.
- **Agrammatism** refers to speech that is lacking in key grammatical information. (In English this information is mainly encoded by grammatical words, such as prepositions and auxiliary verbs).

When aphasia was first identified by Broca and Wernicke, there was a distinction between people who could no longer speak fluently and exhibited word-finding difficulties, and perhaps agrammatism to some degree (Broca's aphasia), and those whose speech was quite fluent, but showed significant issues with paraphrasia and difficulties with comprehension (Wernicke's aphasia).

The following illustrations of Broca's and Wernicke's aphasia were recounted by Howard Gardner in his book *The Shattered Mind: The Person after Brain Damage*. Gardner is the interlocutor in the first example:

(16.5) Broca's aphasia

The patient here is Mr Ford, a 39-year-old Coast Guard radio operator:

'I am a sig … no … man … uh, well, … again.' These words were emitted slowly, and with great effort. The sounds were not clearly articulated; each syllable as uttered harshly, explosively, in a throaty voice. With practice, it was possible to understand him, but at first I encountered considerable difficulty in this. 'Let me help you,' I interjected. 'You were a signal … ' 'A sig-nal man … right,' Ford completed my phrase triumphantly. […] 'Why are you in the hospital, Mr Ford?' Ford looked at me strangely, as if to say, Isn't it patently obvious? He pointed to his paralyzed arm and said, 'Arm no good,' then to his mouth and said, 'Speech … can't say … talk, you see.'

(16.6) Wernicke's aphasia

The patient here is a 72-year-old retired butcher:

> ‘Boy, I’m sweating, I’m awful nervous, you know, once in a while I get caught up, I can’t get caught up, I can’t mention the tarripoi, a month ago, quite a little, I’ve done a lot well, I impose a lot, while, on the other hand, you know what I mean, I have to run around, look it over, trebbin and all that sort of stuff. Oh sure, go ahead, any old think you want. If I could I would. Oh, I’m taking the word the wrong way … ’

Communication problems associated with right-hemisphere brain damage impact a range of functions related at least indirectly to language, including visual spatial awareness and executive functioning, and can directly impact individuals’ ability to communicate effectively. People who have experienced injury to the right side of the brain may have difficulty initiating and maintaining appropriate conversational topics, and managing effective use of eye contact and turn-taking.

The research on language and the brain is clearly vast, and it is constantly changing in response to new technologies and new discoveries in such areas as brain mapping and neuroplasticity (even in adulthood, the brain is not as set in its ways as once believed). We have barely touched the surface here. To conclude, we recommend you visit an extraordinarily beautiful interactive brain map called *The Brain Dictionary*, which scientists have been able to create using functional magnetic resonance imaging (fMRI). These scientists measured the brain activity of seven people as they listened to two hours of stories. The brain map this generated showed activity across the brain (so spread across both hemispheres). It also showed how the words in our brain’s dictionary cluster together in meaningful categories (so organised more like a thesaurus than a dictionary). You can see this dictionary in our heads at work in a *Nature* video on YouTube, and you can explore the brain model for yourself at the Gallantlab website.

Further reading

An excellent introduction to research on language, culture and the mind is Deutscher (2010). He makes reference to a range of other authors who have worked in this field and provides the full story on research into the relationship between colour, language and thought; one of the earlier works is Mauthner (1982); see also Tan et al. (2008). Keith Chen’s controversial TED Talk on the putative link between our language and our ability to save money can be found at https://www.ted.com/talks/keith_chen_could_your_language_affect_your_ability_to_save_money. Useful introductory texts on psycholinguistics include Aitchison (2007), Field (2005) and Warren (2013). Details about brain composition were drawn from McIlwain and Bachelard (1985), and Swanson (1995); for research on lateralisation, language and handedness, see Knecht et al. (2001). For a full account of Phineas Gage, and the neurological aspects of swearing and other taboo-breaking behavior more generally, see Finkelstein (2019); we also recommend Harvard University's Harvard Brain Tour website (http://braintour.harvard.edu), which showcases everything there is to know about our changing understanding of the brain. Publications by Jay (for example, 2000) also provide summaries of the linguistic, psychological, physiological and neurological aspects of taboo language; see also Chapter 10 of Allan and Burridge (2006). Pinker (2007b) offers an account of swearing from the perspective of a psycholinguist. The benefits of bilingualism described in this chapter were drawn from Swanson (1995). Dronkers et al. (2007) describe the findings from conducting MRI studies on two of Broca's original samples, and reports of patients who produce speech after damage to Broca's area can be found in Plaza et al. (2009). Information on different research methods in psycholinguistics was drawn from Garrod (2006). The original experiments on colour naming by Stroop were published in 1935. The research on swearing is from MacKay et al. (2004). The cross-modal priming research mentioned here is from Swinney (1979). The data on word recognition timing are from Indefrey and Levelt (2004). See Möller et al. (2007) on slips of the brain and tongue. Material on speech production follows Warren (2013). The material on the Possible Word Constraint is drawn from Norris et al. (2001). One of the first studies to address issues relating to research on cognition in bilinguals was Peal and Lambert (1962). Adesope et al. (2010) provide a review of a number of studies on the benefits of bilingualism. See Ayçiçegi-Dinn and Caldwell-Harris (2009) on interactions between specific languages and emotion in bilinguals. A useful resource for readers who are interested in thinking further about how to support children being educated in a second language is Barker (2000). We also recommend you visit some of the work on the mental representation of bilingualism by Riehl (for example 2010). There are some impressive images of the denser grey matter in the bilingual brain to be found on the internet (see, for example, Mechelli et al., 2004). Material on language disability has been drawn primarily from Damico, Müller and Ball (2010). Anyone considering a career in education or speech pathology is particularly recommended to read Chapter 1 of that volume. Douglas and Stirling (for example, 2012, 2016) have published widely on play, storytelling and pretence in both neurotypical children and in children with Autism Spectrum Disorder. See also Johnston and Schembri (2007: 17–18) on the impact of aphasia on sign language users. Examples of Broca's and Wernicke's aphasia come from pages 60–1 and 68 respectively of Gardner (1974). The *Nature* article behind the extraordinary brain map is Huth et al. (2016). For more on the dangers of priming in criminal trials, see Fraser (2018); the quotation comes from Fraser's account of the internet phenomenon 'Yanny and Laurel' (in case you missed this meme – YouTuber Cloe Feldman posted a four-second audio clip on Twitter on 15 May 2018 and asked her followers whether they heard a voice saying Yanny or Laurel – the debate then spread, as these things can do, like wildfire: see https://forensictranscription.com.au/laurel-yanny.

Exercises and discussion points

1 Language and culture

Part A: Parameters of kin systems

Review Tables 16.1–16.4 and identify all the different parameters that contribute to the kinship systems involved (for example, birth order). How do different combinations of parameters affect the system overall? How widely shared is each of the parameters you have identified in the languages shown in this chapter? Do you know of systems in other languages with similar or different parameters?

Part B: Gurindji kinship

Some of the basics of the Gurindji kinship system are set out in Table 16.1. This table only hints at the complexity of the Gurindji system, which is based on a set of eight skin groups. These groups are shown in the following table. All the J-initial words refer to men's skin names while all the N-initial words refer to women's skin names.

Male skin names	Female skin names
Jampin	Nampin

Janama	Nanagu
Japarta	Nimarra
Jukurtayi	Namija
Jurlama	Nawurla
Jangari	Nangari
Jangala	Nangala
Japalyi	Nalyirri

Figure 16.11 shows how relationships are organised within the system. Children are assigned to a skin group depending on the skin group of their mother. Children are in the next pair of names down from their mother (so if a woman were Nalyirri, her children would be Jurlama/Nawurla). Solid lines indicate preferred marriage partners and broken lines indicate alternates.

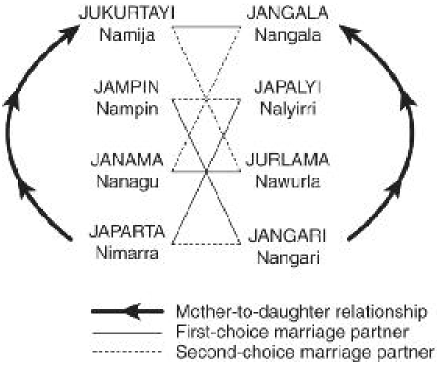


Figure 16.11 Relationship organisation of Gurindji kinship system

(a) Using the diagram, determine who your daughter could marry if you were a Nangari woman.

(b) Now determine who your son could marry if you were a Jampin man.

2 Word retrieval

Versions of the following text circulate on the internet. Try to read the text to figure out what it means. What cognitive strategies support you in reading the text?

> Aocdcrnig to rseecrah at Cmabrigde Uinervtisy, it dseno’t mttaer in waht oderr the lterets in a wrod are, the olny irpoamtnt tihng is taht the frsit and lsat ltteer be in the rhgit pclae. The rset can be a taotl mses and you can sitll raed it whoutit a pboerlm.

3 Tongue-twisters

Using the following tongue-twisters (or with three of your own selection), observe what sorts of errors are introduced when people are asked to say the sentences quickly and repeatedly (allow your friends, family or classmates to read the tongue-twister silently before they start).

(i) Pink lorry, yellow lorry.

(ii) How can a clam cram in a clean cream can?

(iii) Santa's short suit shrunk.

Things to consider:

(a) Which sounds are most likely to cause errors?

(b) Where do they occur?

(c) Is stress a relevant parameter?

4 Sentence comprehension

Consider the following garden path sentences. For each sentence, explain how difficulties arise for the reader. What structures contribute to these difficulties? Are there any examples where reading the sentence out loud limits the garden path effect in comparison with reading? If so, how is this happening?

(a) When Chelsie reads books get broken.

(b) The fast run the course twice.

(c) The first pet that I had loved carrots.

(d) The car raced past the finish line crashed.

(e) I persuaded her sisters are above reproach.

5 Language disability

How are metaphors, idioms and proverbs different from each other? Which of the language disabilities described in this chapter are likely to impair someone's ability to understand or use these figures of speech? How would this effect be caused?

Research project

Planning and performing speech

In this project, you have the opportunity to explore differences between formal and informal speech. Your first task is to obtain or find recordings of two different speech events – one that is formal but unscripted (for example, a speech, lecture or more formal style of radio interview) and one that is not (for example, you could record dinnertime conversation with friends or family and select a longer passage from one person). If you are doing the recording yourself, make sure you obtain permission from the people you record. If you use material that has been broadcast, you do not need to obtain permission to use the material in this way.

In each of your recordings, select a stretch of 90 seconds and transcribe it using the conventions we describe in Chapter 13. Ensure that you mark each transcription to show stressed (loud) syllables (use underlining) and pauses (indicate how long they last); also time each sentence. Answer the following questions as part of your write-up. Your task is to relate your findings to the description of language production in this chapter.

1 How many pauses are there in each transcription? How many of the pauses are planning pauses and how many are rhetorical pauses? How long are the pauses of each type?

2 How many times does the speaker in each text signal hesitation using fillers such as *um* and *er*?

3 Where are stressed syllables found in each text?

4 How much phonological reduction can you identify?

5 How many syllables per second are produced in each transcription?

For more activities, please go to www.cambridge.edu.au/academic/linguistics.

17

Language acquisition

◆

With each newly learned language you acquire a new soul.

Slovakian proverb

17.0 Introduction

The acquisition of something as mind-bogglingly complex as a human language must surely be up there among our greatest achievements. It also represents the highly unusual situation where tiny toddlers actually learn better than adults – children acquire any language or languages with astonishing facility, and they appear to do this without any special instruction. In his book *All about Language*, Barry Blake expresses it nicely:

> From the point of view of understanding language, the difficulties facing the second language learner are of interest largely because they highlight the almost miraculous ease with which children acquire language when they are too young to learn any other area of study. Music provides an interesting comparison. At one extreme there is Mozart composing at five, at the other extreme there are people like me who have difficulty singing the national anthem in tune. With language, every child is a Mozart.

Research in the area of language acquisition is concerned with questions about how these little Mozarts develop skills in using language. This research is a great example of an area in which linguistics makes a contribution, interwoven with fields of research such as anthropology, cognitive science, education, neuroscience and psychology. It is concerned with the acquisition of both first and second languages by monolingual and bilingual speakers, and by speakers showing both typical and atypical pathways of acquisition.

Across these different populations, language acquisition research is concerned with how all aspects of language are acquired: from sound segments through word choice to increasingly complex grammatical structures, the ability to interpret meanings that are only implied, or to construct extended narratives in culturally recognised structures.

Findings from research on language acquisition can be applied in domains like education and in a range of disciplines concerned with early childhood development more broadly. Debates in this area also have connections with our understanding of the evolution of language, since any account of language acquisition also has to map coherently onto our understandings of how language in the brain works more generally (see Chapter 16).

17.1 Approaches to language acquisition research

> Language is an art, like brewing or baking … It certainly is not a true instinct, as every language has to be learnt. It differs widely from all ordinary arts, for man has an instinctive tendency to speak … whilst no child has an instinctive tendency to brew [or] bake …
>
> Charles Darwin (1871), ‘Mental Powers’, in *The Descent of Man*

There are two basic approaches to research in language acquisition, which we will call the observational and logical approaches. The **observational approach** focuses on what we can observe about how children learn language and theorises on the basis of these observations. It assumes that children take the language that is used around them as **input** and learn from it. Research from this perspective is sometimes called **social constructionist** or **emergentist,** since it is concerned with the emerging models of language that children construct for themselves based on their social interactions. Also, within this broad area, some researchers are particularly concerned with the way that children learn to interpret the language used around them. This involves knowledge about more than sounds, words and grammar, since the child must also make inferences about intended meanings of others by considering how language is structured and used in particular circumstances. Even very young children with limited resources in terms of vocabulary and sentence types can be highly effective conversationalists. This type of research has a focus on socio-pragmatic aspects of language in use and on the communicative functions involved. The observational approach leads to an interest in the path of language acquisition: What happens? In what order? With what types of variation? With what intentions? And, with what successes? The focus on learning through interaction has sometimes meant that research in this area has been characterised as prioritising ‘nurture’ over ‘nature’.

The **logical approach** seeks to infer the process of language learning from a model of adult language, which is understood to be the destination of each child’s language learning journey, so the focus is on the endpoint of acquisition rather than the path. This approach generally assumes that children are born with innate knowledge about the structure of language. As a reflection of this assumption, these approaches are also commonly called **nativist** approaches. They generally assume that children’s innate knowledge revolves around a set of rules for language that are associated with a corresponding set of linguistic categories. This focus on rules and categories is termed a **generativist** approach. The abstract rules and categories that are assumed to be in place are known as **universal grammar**, since at birth the innate structures do not concern any particular language and we know that the child will learn the language(s) of whichever community they are raised within. Research developed within this broad framework is focused on describing what is possible in particular languages and in language more generally, asking: What can people say? How does this vary across languages? How are these patterns matched or missed in children’s speech? What mistakes do they make? When? And, why? The focus on the search for evidence of behaviours that are genetically encoded has sometimes meant that research in this area has been characterised as prioritising ‘nature’ over ‘nurture’.

Researchers vary in their willingness to take conciliatory stances in relation to these different approaches. Some researchers acknowledge that at this stage different theories are good at describing and explaining different aspects of language acquisition and are therefore complementary. Others take the position that these two approaches are fundamentally opposed and cannot be reconciled. Nevertheless, there does seem to be general agreement that it is not particularly useful to frame the debate as an opposition between nature and nurture, since language only develops when children have the opportunity to interact with others, while, at the same time, there clearly are many aspects of our brains and bodies constructed through our genetic makeup that facilitate language development in various ways. Most researchers now agree it is always a matter of nature through nurture.

In the following sections, we review the key steps that have been shown to occur as children develop their language skills in the areas of phonology, vocabulary, syntax and pragmatics (as discussed in Chapter 14). Later in the chapter, we also consider some key aspects of literacy development, second-language acquisition and growing up as a bilingual person.

17.2 Language acquisition pathways

It is important to bear in mind how closely language acquisition is associated with brain development. Growth of the brain during childhood is significant. At birth, the brain is about 25 per cent of adult size, increasing to 80 per cent of its mature size during childhood. The structures within the brain also change significantly. A process of pruning neural connections occurs, such that fully one-third of the connections present at birth are lost by adulthood. However, new connections continue to be made, with increasing numbers of synapses per neuron. The connections between neurons are sheathed in a fatty substance called myelin that insulates connections and makes transmission of signals between cells more efficient. This process operates most intensively during the first year of life, but continues into adulthood.

One of the most interesting and exciting things about language acquisition, and child development more generally, is that learning is such a dynamic process. Every new thing the child learns about language or the experience of being a speaker changes the context in which they will learn the next thing. Age is an important factor in language-acquisition processes and you will find that it is mentioned frequently in this chapter. The convention is to represent it by the year followed by the month with a semicolon in between. Thus, four years and three months is represented as 4;3.

In the following sections, we describe the typical pathway that children take in language development. The best-understood pathway is analytical. It begins by building up a repertoire of sounds, then single words, then increasingly long and complex combinations of words. Within the age ranges indicated below, there is a great deal of variation among children, and not all children take the developmental pathway we describe. After all, children are little human beings, not machines, and they aren't following a timetable. Albert Einstein didn't start to talk until he was about three or four; equally, a child who has advanced speech at a very young age isn't necessarily a candidate for the high-IQ Mensa society. There is wide variation in terms of when children reach these milestones and the ages we give here are simply approximations. Although this developmental pathway is common and well documented, it is not the only way children tackle the task of learning language. An alternative method is the **gestalt** approach, whereby children reproduce whole utterance chunks rather than single words (*gestalt* means ‘a whole that can’t be described by the sum of its parts’). This strategy poses significant challenges for researchers, since it is much harder to transcribe and interpret longer strings given the amount of phonological adaptation associated with them. For example, it can be difficult for researchers to sort out whether the child perceives a string of sounds as a unit (i.e. a single piece of language) or has analysed the sounds and then assembled the unit. In practice, it seems likely that children make some use of both approaches, since even children using the analytical strategy occasionally reproduce much longer unanalysed sequences of words.

There are also children whose productive use of language occurs significantly later than it does for others. Within this set of children, some will catch up with their peers by 7;0, as Einstein clearly did. For others, however, the late emergence of language is associated with other language impairments. Researchers do not currently have the means to distinguish reliably during the toddler years between children with specific language impairments (SLI) (see also Chapter 16) and children who are simply late talkers. Diagnosis would be useful for this age group, since children with SLI would be appropriate candidates for therapeutic intervention; however, this would not improve outcomes for otherwise normally developing children in the long run.

A note about ‘input’

The term **input** refers to the language that babies and children hear around them. This especially includes the language that carers use when interacting with their charges, which is known as **child-directed speech (CDS)**. An important assumption that initially motivated the nativist approach to language acquisition was the claim that the language children hear around them is too fragmentary and poorly formed to act as a reasonable resource for a child learning language without built-in supports. How do they acquire a systematic knowledge of the structure of their language when the data to which they are exposed is so impoverished? In particular, in the 1960s, mothers were portrayed as using defective language (what we would call ‘baby talk’) or motherese with their babies and toddlers. Motherese was labelled ‘degenerate’ because researchers assumed it did not provide children with a complete working model of adult language.

Subsequent research showed, however, that the modifications made in CDS are generally quite skilfully attuned to the needs of babies and small children in terms of their preferences and needs. For example, very young babies do not have a well-developed ability to detect changes in pitch or volume and are more sensitive to higher-pitched sounds in general. The sing-song quality of much CDS to babies is consistent with these differences, showing more exaggerated pitch changes along with changes in loudness that invite the attention of the baby. Other aspects of CDS that demonstrate adaptation by adults and older children to the perceived needs of the young child include recasts and expansions of utterances made by the child. A **recast** occurs when the child attempts to say something and part of their utterance is ungrammatical. For example, *I slood down the slide!* might be recast as *Oh oh! You slid down the slide!* An expansion is similar, providing an opportunity to model correct forms, but also to elaborate further on what has been said. An **expansion** for the child utterance *I slood down the slide!* might be *Oh, really?! You slid down the slide! Was it super speedy?* Researchers have also observed that many daily interactions provide opportunities to practise communicative skills with children.

Interestingly, although CDS is widespread in many cultures of the world, there are also well-documented cases of communities in which babies and young children do not receive this type of input because adults and older children do not treat them as **conversational partners** until they can talk back. Input for these children must be more indirect, since they clearly observe many more interactions than they are invited to participate in. Nevertheless, provided these children have secure places in the family structure and are actively incorporated into community life, they learn language successfully – just as other children do. This indicates that there are multiple ways in which children can access adequate input for language learning.

Even so, in those communities where adults enthusiastically engage in conversation with their small children, people are strong in the belief that this sort of behaviour is crucial to the linguistic development of these children. Steven Pinker gives an interesting comparison. The !Kung San in the Kalahari Desert are one of those groups of people who do not speak to their children until they have learned to speak back. Of course, these children learn to talk, just like other children do. On the other hand, the !Kung San believe that children have to be instructed on how to sit, stand and walk. Accordingly, parents carefully go through the routine of piling sand around their infants to ensure they will learn how to sit up. The fact that they do this is proof the procedure works. So, what parent would not engage in this activity? Even today's Western 'helicopter parents', who hover over every aspect of their children's lives, would never dream of training them in this way. This behaviour seems absurd – children learn this stuff on their own. Meanwhile, Western attempts to engage tiny infants in conversation would seem just as absurd to the !

Kung San – why would you speak to children until they've learned the art of conversation?

Janyarrp: Gurindji baby talk

There has been considerable discussion in the research literature about the extent to which special forms of child-directed speech occur across cultures. In the following excerpt, linguists Caroline Jones and Felicity Meakins describe the baby talk register of the Gurindji community as showing a combination of common and culture-specific features:

> Gurindji women held a strong view that baby talk helped children to learn to speak. There was no suggestion of a concern that using baby talk to children would slow their acquisition of language. Gurindji women instead described how they fine-tuned their use of baby talk to the children's improving productions. They also clearly distinguished baby talk from 'double tongue' – that is, speech sound disorders.

As in many other cultures, Gurindji caregivers moderate their use of baby talk as children develop, responding to the increasing abilities of the child. In contrast to the Western cultural context, Jones and Meakins report that the phonological modifications associated with baby talk are used by caregivers until quite late in the childhood of their charges. Gurindji caregivers report a strong focus on supporting comprehension in their use of baby talk – and this is in contrast to the focus on engagement typical in a Western setting. They are also much more focused on the use of baby talk as a model for their children to follow than in Western cultures (compare the discussion on defective input above).

When we listen to someone else speaking, we are engaged in interpreting a continuous stream of sounds. As fluent speakers of and listeners to a language, we have become adept at imposing a structure on the sound stream and identifying meaningful units within it. This is also a significant task for babies and young children in the process of language learning. Words and phrases eventually need to be **segmented** out of larger structures. Moreover, the strings that adults produce regularly involve a range of longer sequences that are sometimes used in highly routinised ways. This regular repetition of conventionalised phrases and idioms gives them the appearance of discrete structural units. In fact, researchers believe that adults, themselves, do not reconstruct these sequences, but rather store and use chunks of them as units.

Wherever they focus their efforts to begin with, children have to develop the knowledge and skills involved in breaking down all types of input, including conventionalised units, into their constituent parts and learning to recombine these in creative ways. For children who begin their careers as speakers of a language with a focus on the gestalt strategy, reproducing extended utterances, a key task in language acquisition is to begin to identify words within the utterance. At the same time, these children must also begin to identify the structural patterns that organise words in relation to one another.

The critical period hypothesis

The **critical period hypothesis** is the idea that there exists in childhood a window of opportunity – or a critical period – during which language must be acquired. This has been linked to brain development. At birth, there is relatively little difference between the functioning of the two brain hemispheres, but around 2;0 this process of lateralisation begins, whereby the right and left hemispheres each take over a dominant role for certain mental functions. As we described in Chapter 16, the relationship between language and lateralisation is an extremely complex one and is the subject of continuing neurolinguistic research. And while the development of this neural one-sidedness would appear to overlap with the main period of time when first-language acquisition takes place, this idea of a critical learning period is still a controversial one. Below is a glimpse of some of the evidence to support its existence.

1 *The different patterns of recovery in brain-damaged adults and children.* Full recovery is seldom achieved if the injury occurs after puberty, but with very young children even the total removal of the left hemisphere doesn't appear to prevent the reacquisition of language.

2 *Socially isolated children without language.* The most famous of these children is Genie (who, tragically, was locked in an attic in Los Angeles from around 2;0 until she was discovered at 13;5). Efforts were made to teach Genie language and, while she made progress in learning vocabulary, she never acquired normal grammar, particularly syntax. It is difficult to know what conclusions to draw from the experience of someone whose

childhood was so traumatic (and Genie may even have had some left-hemisphere brain damage), yet there are other cases of children who grow up in loving families and are without language (for example, Chelsea, whose deafness was only discovered when she was 32;0).

3 *Children with Down syndrome.* These children follow the usual pattern of development, but at a slower rate. There are all sorts of different experiences here, but it seems that if language development is delayed to beyond puberty then typical linguistic skills are never fully acquired.

4 *Learning a second language.* No matter how much exposure you get to a second language, if this exposure occurs after childhood, you will probably never sound like a native speaker. Brain scans support this experience; when acquired in adulthood, second languages are located separately from first languages, but if the learning experience is during the early language acquisition stage, then both first and second languages tend to be found together in the frontal cortical areas of the brain.

Evidence does suggest that there is a window of opportunity in our early lives that is set aside for the acquisition of language (just like some birds never acquire a normal song if they aren't exposed to it early enough). Our brain gets set in its ways, and if linguistic skills are not acquired by that time, then unfortunately these skills will remain deficient.

17.3 Acquisition of phonology

At the same time as children learn to produce the sounds of their language, their heads and mouths grow and develop in several significant ways. For example, a baby's mouth contains sucking pads (fatty deposits in the cheeks that support sucking), rather than teeth for chewing. The tongue is also relatively large, and has restricted movement, being closely connected to movements of the jaw. These features are present into the third year. Around this time, the sucking pads disappear, the space inside the mouth becomes larger and the tongue moves more independently of the jaw. These physical changes continue throughout childhood and adolescence. By 6;0, the skull reaches adult size, but the mandible, tongue and lips continue to grow until late adolescence.

Table 17.1 summarises the pathway of phonetic production for speakers of English. It includes information on consonants, vowels, syllable structure and prosody (intonation) where available. The overall trends shown in the table include a relatively early acquisition of vowels pronounced in isolation as well as oral and nasal stops. Fricatives and liquids are acquired later and children's skills at pronouncing consonant clusters and vowels in polysyllabic words can take several years to be established reliably.

Table 17.1 Ages and stages in acquisition of phonology

Year	**Typical features**
0–1	• **Vocalisations** are produced from very early on, with sounds variously called gurgling or cooing. • **Babbling** begins at around 0;6. • Consonants typically include /m/, /h/, /d/, /n/, /t/, /b/, /k/, /ɡ/, /w/, /s/ (roughly in order of frequency). • Vowels are more frequently low and unrounded. • Prosody is in evidence between 0;10 and 1;0, initially characterised mainly by flat or level pitch contours and variations in loudness.
1–2	• Toddlers begin to utter words sometime between 1;0 and 2;0. • Words have variable pronunciations. **Phonological adaptations** to words include final consonant deletion, cluster reduction, fronting of velars, stopping, gliding and context-sensitive voicing. • Consonants still include /m/, /h/, /d/, /n/, /t/, /b/, /k/, /ɡ/, /w/, /s/. • By around 2;0, vowels pronounced in isolation have expanded to include /i/, /ɪ/, /u/, /ɛ/, /e/, /o/, /ɔ/, /ɑ/, /æ/. • Greater diversity in types of intonation contours available (e.g. high falling and high rising contours now used).
2–3	• Consonants now additionally include /p/, /ŋ/, /j/, /ʒ/, /f/. • Consonant clusters are in evidence though not always consistent with adult forms. • Frequent use of [w] in place of [ɹ] in second position. **Simplification of consonant** clusters still frequent. • Many studies consider that vowel production is generally reasonable accurate by 3;0. • Increasingly complex syllable shapes (monosyllabic words contain V, CV, VC, CVC, CCVC, CVCC, CCVCC, CCVCCC and CCCVC syllables, and polysyllabic words contain V, CV, VC, CVC and CCVC syllables). • Tendency to delete unstressed syllables in non-final position and unstressed syllables with reduced vowels.
3–4	• Accuracy for vowels is above 97% in monosyllabic words and above 88% in polysyllabic words. • Liquids may not be acquired until close to 5;0 and lisps are common

	until 7;0, but in other respects the consonant inventory has been acquired. • Consonant clusters continuing to develop; simplification is less frequent though clusters still show considerable adaptation (e.g. [w] in place of [ɹ]). • Deletion of unstressed syllables now less common.
4–5	• Liquids (/l, r/), consonant clusters and vowels in polysyllabic words continue to be a challenge for children in this age group. • Children also show some awareness of syllable segmentation and rhyme.
5–6	• Control of liquids, but continued variation in the pronunciation of vowels in polysyllabic words. • Types of phonological awareness include syllable segmentation, rhyme, alliteration, phoneme isolation and beginning letter knowledge.
6 +	• Improved accuracy with consonant clusters; these are generally stable by 8;0. • Vowels in polysyllabic words are generally accurate after 7;0. • Children show syllable segmentation, rhyme and alliteration awareness, phoneme isolation and increasing letter knowledge.

The fis phenomenon

Children's production lags behind their perception – in other words, they may hear the difference between (sequences of) sounds and have adult-like mental representations of them long before they can actually produce the different sounds. For instance, they might perceive the difference between a *goat* and a *coat* even though they pronounce both words as [gəʉt]. And they will be able to distinguish *goats* and *coats* in picture books, even though they pronounce them the same. Consider the following now-famous exchange between an adult and child.

Child: Look at my fis!

Adult: Is this your fis?

Child: No, my fis!

Adult: Your fis?

Child: No!

Adult: Your fish?

Child: Yes, my fis!

What has now been dubbed 'the fis phenomenon' comes from an incident described by researchers Jean Berko Gleason and Roger Brown in 1960. This child here is referring to the plastic inflatable fish as [fɪs], but when a parent asks if it is a ‘fis’, the child says, ‘No, my fis’. When asked if it is a ‘fish’, the child agrees. Clearly, this child can understand words without being able to pronounce them – the perception of the phonemes comes before the child’s ability to produce these phonemes. This mismatch between awareness and production explains why children can get so irritated when people try to imitate their baby pronunciations.

17.4 Acquisition of vocabulary

Children show signs of recognising words and phrases from around 0;6. For example, by this age, babies may turn their heads expectantly towards someone who says their name. By 1;6, toddlers are pointing to body parts in response to prompts such as ‘Where’s your nose?’ Table 17.2 indicates typical ages and stages in vocabulary acquisition up to 3;0. Productive vocabulary size refers to words that the child produces rather than words that are understood. Toddlers typically understand many more words than they actively use. In the early stages, it is also important to note that some of the words used by toddlers may not map clearly onto an identifiable adult word. For example, the words *glick* or *nimnam* might be used to refer to ‘milk’.

Table 17.2 Ages and stages in vocabulary acquisition

Age	**Stage**	**Productive vocabulary size**
0;6–1;0	Early commencement of speech	May say a small number of different words
1;0–1;6	Typical commencement of speech	Typically able to say a small number of different words
1;6–2;0	Expanding vocabulary (two- and three-word utterances)	Around 50 words
2;0–3;0	Simple phrases and sentences (words carrying plural and past tense marking)	Up to 1,500 words

One notable aspect of the way children use words is their tendency to over-extend the meanings of words, especially in the early stages when their overall vocabulary is small. For example, any bird that swims may be labelled a *duck*, any furry pet may be labelled a *dog* (including cats and rabbits). **Over-extension** typically is based on shape and function, but the semantic relations that allow for over-extension can be quite creative. Anything that should not be touched might be *hot*. Children also produce **under-extensions**. In these examples, a word is applied to a restricted set – for example, only the child’s shoes may be covered by their word *sooze*. This appears to be less usual than over-extension, although it may also be that it is simply not as obvious when a child under-utilises a word in this way.

Once again, production lags behind perception. A child might use [mʉː] to refer to the moon, half-eaten cakes, marks left on a window and bananas (which all share a certain shape!), but correctly identify the moon when called upon – and the child would certainly take issue with any adult who described a piece of half-eaten cake as a *moon*! Children may be over-extending words they already know simply in order to communicate. They want to be able to say something about objects whose name they haven't yet discovered, or else they can't remember or can't produce. With time, children will narrow down the application of terms as they learn more and more words.

Acquisition of sign languages

Children who acquire a sign language from parents who are native users of that sign language develop in very similar ways to children who learn spoken languages from their parents. They follow the same pattern of building up their use of signs from one to two signs. Furthermore, the types of signs acquired first are similar to those words chosen by speaking children – mother, father, milk, dog and goodbye are typical of both groups. As signs begin to be used in larger sets, grammatical conventions are learnt. Interestingly, research has shown that when hearing children are raised by a couple that includes one parent who is a native signer and another who uses a spoken language, children acquire both and do not show a preference for one or the other.

17.5 Acquisition of grammar

In the 1970s, Roger Brown undertook a detailed study of the acquisition of grammar and discovered that children acquire functional words and morphemes in a particular sequence. Table 17.3 shows the typical order for the first ten grammatical features of English.

Table 17.3 Typical order children acquire English grammatical forms

1	present progressive (*singing*)
2	prepositions *in* and *on* (*in bed; on blanket*)
3	plural inflection (*houses*)
4	irregular past tense verb forms (*went, broke, brought*)
5	possessive inflection (*Daddy's ball*)

6 Copular (or linking) verb *be* (*is, am, are*)
7 Articles (*a, an, the*)
8 Regular past tense inflection (*jumped, goed, breaked, bringed*)
9 Regular present tense inflection (*Mamma sings*)
10 Irregular present tense forms (*has, does*).

If you compare stages (4) and (8), it is tempting to think that children appear to be taking a backwards step. At stage (4), they are producing correct forms, such as *broke*, and then later on they produce forms like *breaked* – what's going on? It may not look like it, but this is progress and it is beautiful evidence that children's speech is rule-governed. There is a very good reason why they trot out grammatical irregularities such as *broke* and *went* so early on. Recall that in Chapter 11 we described how the highly irregular forms in language are typically the common everyday words, and children are picking up on them early because they hear them so frequently. However, at this early stage they are not connecting these verbs with their root forms *break* and *go*, but have them as separate lexical entries. They then acquire the correct rule for forming past tense (add *-ed*). And, in the same spirit as their semantic development, they over-generalise and apply the ending to all verbs, including the irregular ones. They must then learn that there are exceptions. So, they reacquire irregular forms such as *broke* and *went*. Now, much older and wiser, they connect (obviously at an unconscious level) these irregular forms with the basic verbs *break* and *go*.

Analogy at work: Wugs that bing

Throughout this book, we have seen examples of analogy at work; for example, in the creation of new words by backformation (Chapter 4) and the regularising of patterns that happens in language change (Chapter 11). Analogy is also at work – and very impressively so – in child language acquisition.

Psycholinguist Jean Berko Gleason (of the fis phenomenon) is probably best known for having created the Wug test. For this test, one of the activities involved showing very young children a picture of a bizarre-looking creature, which they were told was a 'Wug' (see Figure 17.1 – the original Wugs were much cuter that the ones we have been able to reproduce here, but you can see plenty of images of them online).

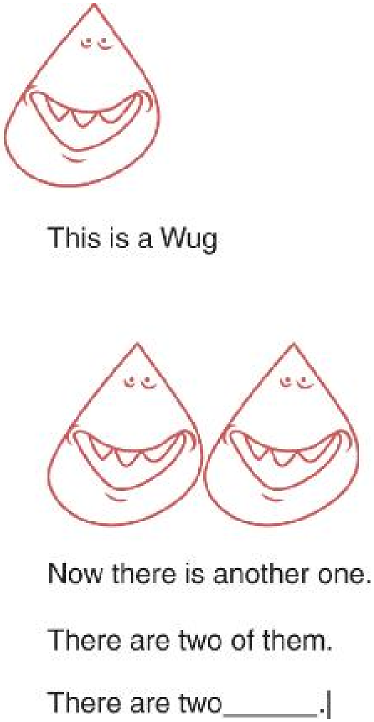

This is a Wug
Now there is another one.
There are two of them.
There are two______.|

Figure 17.1 The Wug test (1958)

They were then shown a picture of two of these creatures and asked, ‘Now in this picture, there are two of them. There are two … ’. Consistently, children contributed the form [wɐgz]. In order to appreciate how interesting this is, we perhaps need to remind you how tricky the English rule of plural formation is.

Consider the following nouns and note how the final sound of each stem triggers a different plural ending.

- *bat* [bæt] – *bats* [bæts]
- *bag* [bæg] – *bags* [bægz]
- *batch* [bætʃ] – *batches* [bætʃəz]

When confronted with pictures of Wugs and other imaginary creatures (perhaps a Zib and a Gotch), young children are able to trot out the correct plural forms (Zibs [zɪbz] and Gotches [gɔtʃəz]). Clearly, they do not simply memorise each plural word as they learn it, but they have internalised a rule of plural formation, which they can even apply to words they have never heard before. This is exactly the kind of reasoning that goes on in analogy.

bug:bugs = *Wug:*x solve x
x = *Wugs*

This is important evidence showing that children look for patterns in the language and don't simply imitate others. Recall the comments we made after Table 17.3 – when children acquire irregular (but correct) forms like *broke* and *brought,* but then start producing regular (but incorrect) forms like *breaked* and *bringed,* this is actually progress (despite what their parents might think!). Similarly, children who can trot out plural forms like *mice* and *sheep* look like they're going backwards when they start saying *mouses* and *sheeps.* But, it's simply that they learn these plural words wholesale. It is only when they start producing forms like *mouses* and *sheeps* that we have the evidence that they have acquired the plural rule. They then have to relearn forms like *mice* and *sheep* as exceptions to the rule.

Berko Gleason also tested other morphological patterns (such as verb tense suffixes), and again children showed that they could come up with the right forms for nonsense words never encountered before. 'This is a man who knows how to *bing* [= man standing on ceiling]. He is *binging*. He did the same thing yesterday. Yesterday he … ?' And the children supplied *binged*!

Putting words together

The earlier section on the acquisition of words and meanings indicated that as children learn an increasingly large vocabulary, an associated task is to combine words into meaningful sequences. This process begins sometime between 1;6 and 2;0 for most children as they begin to make use of two-word utterances. Researchers observing the language produced by children have noted how their use of increasingly complex strings of words rises steadily over time. This has been measured by calculating the **mean length of utterance (MLU)** that children use at different stages. They have also noted that the relationships between words become increasingly complex and the strings get longer. Along with word order, grammatical words and morphemes are an important means of expressing meaning in English, so when researchers calculate the MLU, they also count inflectional suffixes along with words. They then total the number of morphemes and divide this by the total number of utterances (note that for the purposes of child language, irregular forms like *went* count as one morpheme, while *goed* counts as two). Table 17.4 shows how and when English-speaking children build up their MLU – although, as always, there is a great deal of variation.

Table 17.4 Developmental changes in increasing MLU

[illegible] Characteristics
[illegible] multi-word utterances

2;0–2;6	Grammatical morphemes appear	*hafta*
2;7–3;0	Simple sentences emerge	*Mummy go*
3;1–3;5	Sentence elements appear embedded	*I wanna play* *I saw where Kitty went*
3;5–4;5	More embedding, clauses are joined	*We left and Mummy cried* *I think we gotta go*

Acquisition studies of other languages show there are commonalities in the way word order patterns and grammatical markers are learnt, and these are clearly linked to conceptual development. Children show the same sorts of over-generalisation errors, and their skills in linguistic comprehension generally precede their linguistic production. There is, however, cross-linguistic variation that has to do with typological differences, such as grammatical complexity of forms and structures. Speakers of highly inflecting languages, for example, often show precocious acquisition of aspects of morphology compared with that reported for English-speaking children (whose language has less complex grammatical morphology).

The one-word stage

During this stage, children use individual words, often with idiosyncratic meanings and adapted phonology, to communicate with others. Some of these 'words' may be unanalysed phrases like *wasat* 'What's that?' This stage is in evidence from somewhere between 0;10 and 1;8. By leveraging off the shared communicative context and making use of other cues such as intonation and pointing, even a single word can be used in a variety of ways. A few examples are shown in Table 17.5.

Table 17.5 The one-word stage

Communicative act	**Example**	**Extended meaning**
Identification	*sock* [[illegible]]	That's my sock.
Request	*carry* [[illegible]]	Please carry me around.
Comment	*hat* [[illegible]]	I'm putting on my hat like you're telling me to.
Suggestion	*sit down* [[illegible]]	

The two-word stage

Sometime between 1;6 and 2;0, small children begin linking words, and there is a definite period of time in which they are combined into pairs. Usually this happens after they have acquired around 50 words. There are many different ways that words can be related to each other to express more complex meanings and relationships, and since we are missing the grammatical bits and pieces that help us interpret how these words are related, interpretation will depend on context. The phrase *Dadda chair* might express possession ('This is Dadda's chair'); it might be a request for action ('put me in Dadda's chair'); or it might perhaps be a statement of location ('I am in Dadda's chair' or 'Dadda is sitting in his chair'). But, maybe this is attributing too much meaning to these utterances, and really all the child is saying here is: 'I'm talking about Dadda and a chair is involved in some way'! However, rather than being random combinations, research confirms that these sequences are carefully structured. Location phrases are always object/person + location (for example, *Mamma bed; fly chair*). As the selection of strategies shown in Table 17.6 illustrates, the ordering of grammatical subjects and objects conforms to that found in the adult grammar.

Table 17.6 The two-word stage

Semantic relation	**Example**	**Interpretation**
Agent + action	*Daddy go*	Daddy has gone.
Action + object	*pat dog*	I'm patting the dog.
Agent + object	*Mummy bath*	Mummy is filling the bath.
Action + location	*go bath*	I'm going to have a bath.
Entity + location	*dolly bed*	The dolly is on the bed.
Possessor + possession (object)	*Sara boat*	This is Sara's boat.
Entity + attribute	*kitchen hot*	The kitchen is hot.
Demonstrative + entity	*that car*	You have that car not the other one.

We noted above that the measures of MLU include counts of grammatical morphemes. Toddlers begin to use plural marking sometime around 2;0 and, as we saw above, occasionally misapply the regular plural suffix *-(e)s* to irregular nouns (*foots, mouses*). Other high-frequency morphemes that are tackled during this period include the regular past tense marker *-ed*, which emerges from around 2;1 but takes until around 3;4 to master. As with the plural forms, children often go through a stage of double-marking of tense (for example, *didn't caught* where past is signalled twice). Clearly, the children have their own system at each of these stages. They are not simply imperfect imitations of adult speech. The possessive marker *'s* emerges at around 2;4, but takes until 3;4 to master. The third-person singular agreement marker *-s* shows a similar developmental pattern.

Telegraphic speech and beyond

The name for this stage perhaps needs some explanation. The telegraph was a device for transmitting and receiving messages over long distances; the telegraph message was known as a telegram. You may not remember the telegram, but, as we discussed in Chapter 4, they required a clipped way of writing that tried to pack as much information into the smallest possible number of words or characters, so grammatical words were typically missing or reduced in some way (e.g. *back 6 pm* 'I'm getting back at 6 pm); these days, the closest thing to a telegram would be an SMS text or a tweet.

This telegraphic stage starts at around 2;0 to 2;6, when children jump from the two-word stage to the multi-word stage. The most common three-word utterances tend to involve agent + action + object (for example, *me want book*) and agent + action + location (such as *Dooni go school*). Toddlers' speech is described as having a telegraphic quality because it is comprised of words from the major word classes (nouns, verbs and adjectives) without much grammatical support (prepositions, auxiliary verbs and inflectional markers). Then, as grammatical words and inflectional forms are used with greater regularity, children's speech becomes more adultlike. Prepositions begin to emerge relatively early, along with pronouns, although pronouns take much longer for young children to sort out.

Research has also been conducted on a number of specific sentence types in English and fairly consistent developmental pathways are shown in Tables 17.7 and 17.8 for the acquisition of questions and negatives, both of which require the use of auxiliary verbs.

Table 17.7 Developmental pathway for questions

Stage	Strategy	Examples
1	Either add a wh-word at the start of the utterance or use rising intonation.	*Where Davey?* *Help?*
2	Use same strategies but with more complex utterances.	*When Davey come back?* *Mummy help me?*
3	Invert the subject and verb and use auxiliaries.	*When will Davey come back?* *Can you help me, Mummy?*

Table 17.8 Developmental pathway for negation

Stage	Strategy	Examples
1	Add *no* or *not* to the start of the utterance.	*Not go* *No bath*
2	Start using *don't* and *can't*; place *no* and *not* in front of the verb.	*Davey don't go* *Don't want bath*
3	Incorporate auxiliary verbs and the negator (now between the auxiliary and the verb).	*Davey can't go* *Sara don't want bath*

From very early on, children create sentences made up of more than one clause by simply jamming clauses together (*She put a bandage on my foot; maked it feel better*), but at around 2;5 they start to conjoin them with *and* (*She put a bandage on my foot and maked it feel better*). Other developments around this time include the emergence of complex sentences – first with *if*, then *because* and later *when* and *but*.

It is not until about 3;0 that children produce complex sentences with relative clauses, although they are without the relative words (such as *who, which* and *that*). Typically, these clauses appear at the end of the sentence, immediately after the noun phrase they modify: *You write some words me know* 'You write some words that I know'. Children usually don't acquire the full knowledge of this complex construction until they are of school age. Relative clauses that appear in the middle of sentences (*The puppy that is eating the bone is hungry*) are more complex to process. They are rare in adult speech and appear quite late in children's speech.

In your own time, please!

The following two exchanges between mothers and their children, documented by Cazden and MacNeill respectively, are now famous in the language-acquisition literature. In both, the child repeatedly fails to make the corrections, even though the parent is providing the right form each time. We feel particularly sorry for the second child, whose message appears to have been lost in the excitement of data collection.

How to hold a rabbit

Child: My teacher holded the baby rabbits and we patted them.

Mother: Did you say your teacher held the baby rabbits?

Child: Yes.

Mother: What did you say she did?

Child: She holded the baby rabbits and we patted them.

Mother: Did you say she held them tightly?

Child: No, she holded them loosely.

Lonely days

Child: Nobody don't like me.

Mother: No, say 'Nobody likes me'.

Child: Nobody don’t like me.

[Last instruction and response is repeated eight times.]

Mother: Now, listen carefully, say ‘Nobody likes me’.

Child: Oh, nobody don’t like me.

We have already seen plenty of evidence that a child’s linguistic awareness often outstrips their production capabilities, so perhaps we can’t read too much into such exchanges. Nonetheless, even when the parent explicitly instructs the child, their efforts appear unsuccessful. In fact, parents and caregivers are generally far too busy (and far too exhausted) to engage in this sort of linguistic activity too often, so very little of this activity too often, so very little of this kind of reinforcement actually goes on. No one is teaching children language (via direct instruction) – children are discovering it all by themselves. Indeed, as we described earlier, in many societies, children are not considered viable conversational partners until they can talk back.

While explicit instruction might appear futile, there is a link between what children produce and what they hear. Linguistic features that occur frequently in speech are the ones they acquire early on, and those rarely encountered are acquired late. Relative clauses are good examples and so are passives. The passive construction is infrequent in speech, especially when it involves a *by* phrase (*The window was broken by the wind*), and full command of sentences such as these occurs very late (at 8;0–9;0).

17.6 Other things you need to know to be a user of language

The previous sections showed how even a single word utterance can be used to great effect, providing the child has an engaged and responsive conversational partner. In this section, we map out the pathway that children take to communicative competence. Since this is another aspect of development that can be mapped fairly closely into ages and stages, we will use Table 17.9 to summarise the pathway of communicative competence for speakers of English. The table shows that even in the earliest months of life, babies are engaged with others and begin learning how to participate in conversation. By 3;0, most children are effective communicators, although willingness to participate in conversation may at this age be determined by how familiar and well liked their conversational partner happens to be. (This is a pattern that follows some of us into adulthood!)

Table 17.9 Developmental pathway for communicative competence

Year	**Typical features**
0–1	From birth: • caregivers may attribute meaning to verbal and non-verbal behaviour • babies unintentionally provide information through crying and through the way they hold their bodies (e.g. tense or relaxed) By 0;6: • babies begin to use **intentional behaviours** to communicate; this may include waving to say goodbye and pointing to establish joint attention and turning away the head to refuse food • **primitive speech acts** that emerge during this period include calling out, greeting and leave-taking, protesting and requesting actions by others Late in this period: • babies begin to incorporate common intonation patterns into their babbling, producing, for example, question intonation over an uninterpretable string of sounds
1–2	During this period: • toddlers begin talking, using single words for a variety of purposes such as making requests, giving instructions and sharing information • they are able to follow conversations and broadly interpret words they are not yet able to produce Towards the end of this period: • toddlers make spontaneous repairs during conversation, showing that they are monitoring their own speech • they develop skills in adapting their speech to different listeners • they are able to communicate their feelings and respond to the feelings of others
2–3	During this period: • children become increasingly active conversational partners, continuing to build on their gains in terms of pronunciation, vocabulary and grammatical complexity • they are able to follow simple instructions, such as *Put your shoes on*
3–4	During this period: • children can be very persistent questioners, preoccupied with trying to understand the world around them • they can answer questions, listen to responses and hold conversations • they can talk about anticipated events that have not yet occurred • they can use their language to explore their imagination (perhaps by

chatting with imaginary friends)

Acquiring honorifics in Japanese

One socially salient aspect of acquiring Japanese is learning how to use honorifics (in Chapter 10, we described how different levels of politeness permeate Japanese grammar). Furthermore, because a person's social status changes over their lifetime, the use of honorifics is always being relearned. How are children socialised into their use of honorifics? In a paper published in 2013, Burdelski shows how caregivers teach the use of Japanese honorifics in both real-world and role-play situations. Children learn about honorifics from caregivers in three ways: by noticing how they are addressed; through direct instruction from caregivers; and when caregivers speak on behalf of children, modelling appropriate speech. Addressee honorifics are used by children as young as 2;6, while referent honorifics tend to develop later.

One interesting aspect of children's use of addressee honorifics is their deployment in strong directives. The following example from Burdelski shows how a teacher makes her directive stronger by repeating it with an addressee honorific added. Because the honorific denotes more formal, and hence more distant, relationships, this is understood to represent a style shift to a more formal mode of interaction and a withdrawal of emotional support.

1 **Teacher:***Nage-nai.*
throw-negative
'You do not throw'.

2 **Child:**((throws another toy))

3 **Teacher:***Nage-masen.*
throw-negative addressee honorific
'You do not throw'.

Children may also deploy this same technique with each other. The following example shows one girl repeatedly trying to get another girl to stop interfering with her puzzle. After her first attempt fails, she repeats the remark more loudly and slowly. When this has been unsuccessful after two attempts, she then makes use of the addressee honorific, using the same loud and slow prosody, and after one final attempt without the honorific turns to the teacher for assistance.

1 **Girl-1:**Moo hitori de dekiru kara daijobu.
Emphasis one.person by able so fine
'I can do it by myself, so I'm fine (without help)'.

2 Girl-2:((does not withdraw hand)) (0.2)

3 Girl-1:HITORI DE DEKI:RU:::. ((takes piece out of G-2's hand))
one.person by able
'I can do it by myself'.

4 Girl-2:((hand on pile of puzzle pieces)) (0.3)

5 Girl-1:HITORI DE DEKIRU.
one.person by able
'I can do it by myself'.

6 Girl-2:((continues to put hand on pile of puzzle pieces))

7 Girl-1:HITORI DE DEKI-MASU:::.
one.person by able-addressee honorific
'I can do it by myself'.

8 Girl-2:(0.8)

9 Girl-1:HITORI DE DEKIRU:::.
one.person by able
'I can do it by myself'.

10 Girl-1:((gets up from table and makes appeal to teacher))

17.7 Development of literacy

> I struggled through the alphabet as if it had been a bramble bush; getting considerably worried and scratched by every letter.
>
> Pip, in Charles Dickens (1861), *Great Expectations*

Literacy refers to a degree of competence in reading and writing, although exactly what this means varies from one community to another. The ability to read and write simple messages is sometimes called basic literacy. A definition of functional literacy is concerned with the capacity of an individual to participate effectively in communication mediated by written texts and other modes of communication across a range of social contexts (possibly including acts such as map reading, filling in forms, buying a train ticket or using a voice-operated telephone service). Academic literacy refers specifically to the ability of individuals to participate in reading and writing in academic contexts – in these settings, there is often a heavy focus on decontextualised meaning.

We will focus here on the stages involved in learning to read and write that are often associated with the early years of schooling. A number of widely recognised stages are associated with the process of becoming a reader and a writer. These stages reflect both a set of learning tasks and the development of a set of social practices more broadly. There are many different sets of labels for these stages. The labels we use here reflect the major cognitive efforts involved in each stage. We have not attempted to identify these stages with particular ages, since children need access to print materials, and some instruction and support (or at the very least modelling) in order to be able to learn to read. The age at which these opportunities are available varies considerably depending on the literacy practices of the community in which the child lives, and the extent to which literacy coincides with second-dialect or second-language acquisition.

Exploration of a literate identity

In this first stage of development, children explore literacy practices in evidence around them by acting out the mechanics and meanings of reading and writing with friends and carers. Typical examples include:

- ‘reading’ books by reciting remembered passages or by describing the pictures
- ‘writing’ notes, cards, lists and picture captions using a range of scribbles, some of which are recognisable symbols (such as love hearts) and perhaps letters (often the first letter of the child’s name) and reporting the meaning of these notes to others
- relating images from sources such as books and television to their own experiences.

Exploration of text features

In this stage of development, children begin to match some familiar spoken and written words. Their name is one of the first words that most children recognise and learn to write. Environmental print (such as words on a traffic sign) also provides common targets for early reading. During this stage, children come to an understanding that written texts are stable – they do not vary in relation to different readings, but always say the same thing. They also begin to identify relationships between letters and sounds, generally with some instruction from their adult carers.

Early reading

During this stage, with support and encouragement from carers, children become actively engaged in decoding written words. They rely heavily on contextual cues (such as pictures) and on sounding-out strategies to interpret individual words. Because this can be a laborious process, children often struggle to understand the larger meanings of a text during this stage. All their effort is expended on decoding the words as they occur in sequence. English orthography is widely acknowledged as being among the less predictable systems, making the sounding-out strategy problematic in many contexts. For this reason, sight reading – recognising words are whole 'shapes' – is a critical skill in developing fluency.

Reading for meaning

As children become more effective in their reading strategies, and are able to recognise an increasing number of sight words, their reading becomes more fluent and consequently more rewarding to them. They are now able to refocus some of their attention on the meanings of the texts they are reading. This includes the ability to think critically about the text or the situation represented within it. Clear evidence that children are reading for meaning is found when they correct themselves as they read – for example, if they misread one word for another. This shows that they have understood the text they have read to that point and are reflecting on whether the words they are selecting make sense.

Independent reading

As the process of reading becomes increasingly automated, fluency increases dramatically, and children are likely to slow down and resort to early decoding strategies only when they are confronted by new words and difficult texts. As decoding is automated, meaning becomes the primary focus of attention and children are able to think about the meanings within texts in more complex ways – for example, by relating different texts to each other or to personal experience.

Developing writing

Children learn to write through stages that are similar to those involved in reading, with the important difference that writing is a productive process – it allows children to make their own meanings available in the world.

Early writing emerges after exploration of literacy and text features, as described above. It often includes attempts at known letters and the use of spacing. Children refine these practices in the next stage, when they typically add the use of some punctuation. Emergent writing becomes conventional as children improve their use of spacing and punctuation, and are able to produce recognisable letter forms and spellings more consistently. As these skills consolidate, children's proficiency increases so that they are able to communicate independently in print.

Figure 17.2 shows an example of emergent writing in which pictures support the text. The intended message reads '[I] made a mud pie' and shows a person outside (as evidenced by the trees) holding a pie-shaped object. The writer of this text shows a good understanding of text placement and spacing, but does not yet reliably form all letters in English orthography. Spelling is also some way from conventional but shows a good use of sound-to-letter correspondences.

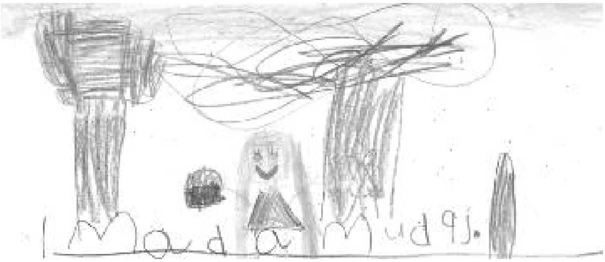


Figure 17.2 Sample of child's emergent writing

Reading development and impairment across cultures

Children display reading impairments in all languages; however, differences in the scripts and in the structure of these languages mean that impairments can manifest themselves quite differently – being dyslexic in Chinese, for example, is not the same as being dyslexic in English. Cammie McBride is a developmental psychologist who researches how children learn to read and write in different languages. She is based at the Chinese University of Hong Kong, an ideal location to conduct this kind of research. Most people in Hong Kong speak both Cantonese and English, meaning children have to acquire two very different writing systems, an alphabetic and a logographic system – the English system that converts letters into sounds versus the Chinese system that converts symbols into meanings and sounds (see

Chapter 8 for more on these systems). McBride's work is all the more important given that early literacy is a field that has had a very Eurocentric focus in terms of the languages that have been researched.

McBride and her research team came up with a task to test young children's awareness of morphemes, and they were among the first to demonstrate the importance of morphological knowledge for reading in Chinese. The test wasn't so different from the Wug test we described earlier, but it concentrated on compound words (compounding is an important structural feature of Chinese languages). So, the child might be asked something like, 'When the sun goes down at night, we call that a *sunset*. What would we call it if instead the moon went down at night?' The child would be expected to answer with something like *moonset*. It seems that children who can give morphologically acceptable answers do better in reading, especially in Chinese and Korean. They also do better in vocabulary knowledge, even when statistically controlling for previous vocabulary knowledge in some longitudinal studies.

Having shown that morphological awareness is critical for Chinese reading, McBride and her team went on to explore the cognitive profiles of children who were at-risk for dyslexia. In this study, they identified several skills, including morphological awareness, that were particularly impaired in children with a familial risk for dyslexia. Subsequently, they found a strong relationship between morphological awareness and reading among children who manifested developmental dyslexia in Chinese two years later. When the condition is undiagnosed, a child's poor performance is often attributed to laziness or stupidity, and the fallout can be disastrous. This early morphological awareness measure predicted subsequent dyslexia in Chinese children – the wonderful result was a standardised screening tool for reading difficulties. Hundreds of paediatricians, developmental behavioural paediatricians, educational psychologists and clinical psychologists have been trained in using this tool, and this means that hundreds of children are now no longer at risk of not finishing their education simply because their dyslexia has gone undetected. Early diagnosis means a significantly rosier future for these otherwise normally developing children.

17.8 Second-language acquisition

> An outsider can master a language as a rider masters his mount; he rarely becomes as one with its undefined, subterranean motion.
>
> George Steiner, 1967, 'K', in *Language and Silence*

The term 'second-language acquisition' (or SLA) is used to cover any language or languages acquired after the first language (or 'mother tongue'); in other words, it also captures those situations where people acquire competence in three or even more languages (so some linguists prefer the label 'additional language acquisition') and the acquisition of languages any time later in life. Given that there are around 200 countries in the world and approximately 7000 languages that somehow have to fit into these countries (ignoring the difficulty of identifying and classifying countries and languages!), monolingual communities will be rare: most children in the world grow up to be 'bilingual' or 'multilingual' – in other words, with an ability in two or more languages.

So, what level of ability do you need to be described as 'bilingual' or 'multilingual'? These labels are by no means clear-cut, and we could spend the rest of this chapter discussing them. Since equal command in two languages is rare, we will assume that bilingualism or multilingualism falls along a continuum – from native-like fluency (for example, someone has learnt two languages simultaneously from infancy) to a fairly limited ability (one of the current authors can read Dutch but can barely speak a word of it).

Research in second-language acquisition has tended to focus on the following three areas.

- How do people acquire second (and subsequent) languages?
- How is second-language learning different from first-language learning?
- How can individual differences in SLA be explained?

More specifically, particular attention has been paid to the influence of instruction and other factors on SLA, including the contribution of a person's first language and the question of whether or not there is a common pathway through SLA similar to the typical phases of development identified for first-language acquisition.

The acquisition of additional languages, however, presents quite a different and more complex range of experiences compared with the acquisition of a first language. For a start, much will depend on the age at which the learning takes place. Some children acquire a second – perhaps even a third – language under very similar circumstances to the first language. Some of you will be in this fortunate position and we look at your experiences in the next section. SLA researchers, however, typically focus on the language used by people acquiring a second language later in life (from the beginning of primary school), with particular attention to questions around the types of errors learners make, their use of formulaic expressions and their acquisition of particular structures (for example, the past tense in English). Just as we saw with children acquiring a first language, there seems to be some disjunction between the structures that learners understand and those they can produce.

As with first-language acquisition and the acquisition of literacy, learning a second language has social consequences since, to varying degrees, it involves people's developing ability to expand their social networks, and access to information and resources in another language.

Developmental patterns in second-language acquisition

Early attempts at speaking a second language in an informal setting are typically characterised by the following:

- preliminary silent period
- formulaic chunks (*I don't know … That one I don't know … I don't know what's this …*)
- telegraphic speech (similar to *no like X* 'I don't like X')
- awareness of the general desirability of grammatical correctness.

As we saw with first-language acquisition, language gets worse before it gets better. This is called U-shaped performance, and although it appears to reflect back-sliding in terms of correctness, it actually reflects steady progress in terms of learning. Table 17.10 illustrates this pattern using the acquisition of past tense marking on strong verbs in English.

Table 17.10 Stages in the acquisition of English strong verbs

Stage	**Example**		**Description**
1	*Yesterday I sing*	telegraphic speech	lacks past tense marking
2	*Yesterday I sang*	irregular forms appear	Instances of use of correct form
3	*Yesterday I singed*	regularised past	overgeneralisation of rule
4	*Yesterday I sanged*	possible use of hybrids	irregular form with regular past suffix
5	*Yesterday I sang*	correct use of strong verbs	target past tense forms

The impact of the first language in second-language acquisition

There are both positive and negative ways in which the first language influences the acquisition of a second language. When the first language and the second language have a feature in common, the process of transfer may have a positive effect on the speaker's performance in the second language. For example, speakers of a language with a strict two-term number marking system for nouns (singular and plural) are less likely to have difficulty in managing the English number marking system than speakers of a language that does not require nominal number marking (resulting in constructions along the lines of *six banana for three dollar*). This unconscious influence of one language while speaking another language is called **interference**.

Where a construction is completely absent from the first language, it is more likely to be avoided in the second language. For example, one of the authors (Tonya) went off to Germany shortly after finishing high school, with six years of school German under her belt and little experience in managing social relationships based on covert social hierarchies. Since German requires the use of special pronouns and verb agreement with social superiors, including most adult strangers, this posed a few challenges, and talking with strangers became problematic since the forms she knew best were over-familiar. For a while, she tried to avoid referring to people directly, with the probable result of sounding self-absorbed instead of over-familiar. It is not clear that the strategy of avoidance resulted in an overall gain!

Conversely, where a structure is used frequently in the first language, it may be over-used in the second language. This is especially noticeable with different types of discourse strategies. For example, there are many cultures that do not require the automatic use of politeness expressions such as ‘please’ and ‘thank you’; however, it can be very difficult for a speaker of English from a middle-class background to stop themselves from using them in a second language, even if it means making something up to fit the purpose.

In earlier studies of bilingualism – particularly where the standard of comparison was monolingual speakers rather than other bilinguals – interference was viewed as being symptomatic of problems. One interesting aspect of interference is that it tends to be more pronounced in stressful situations. Weariness, excitement, and the pressure of being tested, are all known to increase the amount of interference in a bilingual person's speech. Importantly, language learners tend to exhibit more interference when they are interacting with other language learners. This suggests that interference – although it does indeed seem to be largely unconscious – does actually have positive as well as negative meanings for the speaker. For example, when language learners are talking together and showing a lot of interference, they may be signalling group solidarity, another example of accommodation (covered in Chapter 12).

Pragmatic interference

Languages vary in how explicitly they encode information about the speaker's attitude towards the information they are communicating (for example, whether they believe it is true or how they learnt the information). This is also known as epistemic stance. In some languages, this information is coded in the grammar, while in others only more subtle markers associated with prosody and expression are typically used. Using words or morphemes equivalent to, 'I heard that …', 'I have inferred that …' or 'I saw with my own eyes that …', is a way of aligning or distancing the speaker from the message.

In a 2006 study that explored the use of Japanese markers of attitude, Ishida argues that differences in first-language markers of attitude can have an impact on how a person is understood when speaking a second language. When reporting information that has been heard from another person (as opposed to reporting on events experienced directly), for example, there are differences in when English and Japanese speakers will indicate that the information is hearsay. English speakers typically do not bother to indicate hearsay if they consider that the source of the information is reliable, whereas Japanese speakers are likely to use a hearsay marker, particularly if they have just recently heard the news themselves. English speakers who omit hearsay markers in their Japanese can come across as being overly confident about information they receive from others.

Since unexpected patterns used by second-language speakers can trigger negative perceptions in first-language speakers, Ishida argues that it is useful to provide explicit instruction to learners about the typical patterns and forms of encoding epistemic stance.

Success factors in second-language acquisition

Among the success factors that have been identified in acquisition of a second language are the use of diverse learning strategies and aptitude for the tasks involved. Learners differ in the ways that they approach learning. Cognitive strategies involve analysis of language – for example, learning how an utterance can be broken down into units and how various kinds of units can be combined with each other. Meta-cognitive strategies involve conscious thinking about language learning – for instance, planning, monitoring and evaluating learning. Social strategies involve talking about language – for example, asking for clarification. The more of the different strategies learners use, the more successful they will tend to be.

One of the difficulties in talking about aptitude is that successful language learning requires quite different skills, depending on the availability and type of instruction. Aptitude in formal classroom language learning is far more strongly associated with the types of qualities measured in traditional intelligence tests, whereas a person acquiring a second language without formal instruction must make use of interpersonal skills in order to develop opportunities for learning. Aptitude for second-language learning in school correlates broadly with successes in first-language acquisition, including children's use of longer utterances at earlier ages, the size of their vocabulary, skills in comprehension and the level of literacy-oriented behaviour they see at home.

17.9 Growing up bilingual

Studies of children acquiring more than one language from birth have shown that the bilingual child follows similar patterns to monolingual children learning each language. The language that is spoken at home or to which the child has the most exposure will be acquired more rapidly. In the very early stages, children raised bilingually are similar to monolingual children in terms of the age of onset of babbling and the amount of vocalisation they produce. Later, the acquisition of particular grammatical constructions in either language corresponds with their degree of difficulty in each language. For example, prepositions are one of the more complex aspects of English grammar and children tend to over-generalise their use. Even for constructions that apparently are equally difficult, the time of acquisition may differ depending on the frequency with which the child is exposed to the construction.

Acquisition of phonology

We showed earlier that there is an enormous amount of simplification in the phonology of small children because they are not able to articulate all the sounds used in adult language. Children who are acquiring two languages have a greater repertoire of sounds to choose from when using this strategy. They may make use of a sound from either language as a substitution for a difficult sound.

Acquisition of vocabulary

During the single-word stage, monolingual children often over-extend the meanings of words. Bilingual children also begin to use this strategy at around the same time, and they make use of words from either language without regard for the language of the audience. As over-extension is resolved, children will acquire words for equivalent concepts in different languages. They don't resolve over-extension using analytical strategies (if this is called 'paper' in English, what is it called in Spanish and German?) because they don't have the necessary cognitive ability when they are young. The different languages also tend to be used in different domains, associated with different sets of vocabulary.

Some studies have claimed that bilingual children have deficits in vocabulary; however, provided the children who are being compared are matched in methodologically sound ways, it seems that there is in fact no vocabulary deficit – bilingual children have been shown to have the same comprehension of each language as monolingual children and their total production of vocabulary is equal to the productive vocabulary of monolingual children (but spread between the two languages).

Acquisition of grammar

For both monolingual and bilingual speakers, simpler structures are learnt before more complex ones. The kind of language available to bilingual children is also important in determining how they acquire language. Those who are only exposed to the simplest structures obviously cannot make great strides in learning the more complex features of a language.

The morphology of any particular language tends to be acquired in the same order for both monolingual and bilingual speakers. This reflects the importance of morphology for clear communication and the complexity of the systems. The errors made by monolingual children are similar to those made by bilingual children, though bilingual children may take longer to sort out these errors. As in the situation with phonology, bilingual children may opt to use forms from one language in another language if they seem to solve a problem.

Language awareness and differentiation

All children – whether bilingual or monolingual – gradually develop skills in using different varieties of language in different contexts. Their developing skills in this area are part of the larger task of being socialised into their communities. For monolingual children, this involves registers (formal versus informal) and dialects (standard versus non-standard), and so on; for bilingual children, it also includes languages.

Reports on the age at which children show that they are able to distinguish between their two languages vary widely. Many bilingual children are not able to name the distinct language they speak until at least 4;0. How much linguistic knowledge children have and how much they demonstrate may be quite different things. We have already noted that very young children are able to distinguish between the phonological systems of distinct languages, although they cannot name them.

By around 2;0, some children are able to handle their two languages and use them with appropriate people. This is readily observed in families that choose to follow a 'one parent, one language' strategy in raising bilingual children (whereby each parent consistently speaks only one of the two languages to the child – for example, the mother uses German and the father uses English). From this age, children tend to get upset when the languages the parents use are not the ones they use to address the children. Children also dislike or fail to respond to comments in the 'wrong' language from one of their parents. Some children go so far as to translate the words of one parent into the language of the other parent, apparently in the belief that they do not understand each other. Alternatively, in some families the strategy may be to use one language within the home and another outside, or to use one language when the father is present and another when he is absent, and so on. Whatever strategy is used, it is designed to foster a larger process in which children learn that particular languages are appropriate in particular circumstances.

Managing 'one parent, one language' interactions

Nibun and Wigglesworth report on the early ability of one child to differentiate languages used in a family taking the ‘one parent, one language’ approach. This child is reported to have differentiated the language being used from 1;4 and had firmly established differentiation according to their conversational partner by 1;8. The authors note that borrowing between languages was most frequently related to gaps in the lexicon of the language with the smaller vocabulary. This is shown, for example, in the following excerpt, with Japanese forms in bold and German forms in italics.

The child (1;7) opened an animal puzzle book to find puzzle pieces missing.

Child:na:i, nai.
not there.

Father:*was ist das?*
what’s that?

Child:na:i, nai.
not there.

Father:*nichts drin?*
nothing in?

Father:*ist das leer, nh?*
is it empty, huh?

Child:*leer.*
empty.

Father:*leer, keine Tiere drin.*
empty, no animals in.

The benefits of growing up bilingual

As described elsewhere in this book, for many years a widely held belief existed that bilingualism was somehow a bad thing. By adding a considerable cognitive load to children in those vital early years, bilingual language acquisition was somehow thought to be damaging to these children. But research has not borne this out. It turns out there are considerable educational benefits of bilingualism, some of which we touched on in Chapter 16.

- *Flexible thinking*. Bilingual children understand better how language works and are more able to differentiate form from content or meaning – something that is crucial to our everyday thinking. This is a good basis for future cognitive development, especially when it comes to adaptable thinking (for example, the skills needed to meet new challenges, shift strategies and adjust to new situations).

- *Bilingualism and reading readiness*. When you are exposed to a new language, it teaches you about the nature of language and languages, and as literacy experts show, this is precisely the sort of knowledge that literate people need to develop. (Note that this benefit is not restricted to children growing up bilingually – it also applies to, say, very young children who are participating in a primary-school language program, such as Japanese lessons.)
- *Linguistic awareness*. Bilingual children are better able to judge the grammaticality of sentences – they can understand grammatical rules and detect word boundaries more successfully than monolingual children.

There are also benefits to a bilingual upbringing outside the education arena. The phenomenon of globalisation – particularly issues to do with trade, security, tourism and business – is making language a very marketable commodity. From a worldwide perspective, ‘monolingual’ communities, where only one language is used, are rare. As described in Chapter 12, most children in the world grow up learning at least two languages. Australia has always had a multilingual and multicultural population, but it is also a country where English monolingualism has considerable clout. The country is dominated by English – its national language and its lingua franca are English, and English language and English-medium education are among its major industries. Yet, if Australia (or any other country in a similar position) is to compete in a world that is becoming increasingly multilingual, it needs to be taking advantage of its rich linguistic resources.

Further reading

The Mozart quotation comes from page 250 of Blake (2008), and the example of language acquisition by the !Kung San in the Kalahari Desert from page 40 of Pinker (2007a). We have borrowed the labels ‘observational’ and ‘logical’ as approaches to language acquisition research from Foster-Cohen (1999). See also the discussion in Ambridge and Lieven (2011) for contrasting views. Research on language segmentation has been drawn from Peters (1983). Smith

(2002) summarises the various pieces of evidence for the critical learning hypothesis. Jones and Meakins (2013) discuss the Gurindji baby talk register (the quote is from page 188). Tables 17.1–17.3, which describe the development of English phonetic production and the development of increasing mean length of utterances, are adapted from McLeod and Bleile (2003). See Johnston and Schembri (2007: 18–19) for more information about sign language acquisition. Research on acquisition of morphology and making use of the mean length of utterance has developed out of work by Brown (1973). Berko (1958) was where Wugs first made their appearance. The fis phenomenon was described in Berko and Brown (1960: 517–57), and the two examples we provided to show that children don't respond to tuition and correction came from Cazden (1972: 92) and McNeill (1966: 69); see also Berko and Brown (1960). The report on how caregivers teach the use of Japanese honorifics can be found in Burdelski (2013) (the example we quote is from page 261). Japanese markers of attitude are discussed in Ishida (2006). Nibun and Wigglesworth (2014) describe an example of trilingual language acquisition (the example we use is from pages 87–8). In this chapter, we concentrated on English, but for a wonderful survey on language acquisition across a wide range of languages, we recommend Slobin (1985–97). Siegel (2012) is a valuable resource on second-dialect acquisition and Baker (2000) is an accessible starting point for exploring educational issues associated with growing up bilingually. Various myths associated with multilingualism are examined in Tokuhama-Espinosa (2003). Liddicoat (c. 2000–01) describes the literacy interdependence between two (or more) languages regarding the associated benefits of bilingualism. For more on the research into reading development and impairment across cultures, see McBride (2015). Testing young children's awareness of morphemes is discussed in McBride (2016). Another very interesting strand of research within language acquisition explores the way that people commonly understand their children's language learning: see, for example, Niedzielski and Preston (2000, Ch. 4). To help you negotiate the muddle of methodologies, fashions, ideologies and political convictions that characterises the world of multilingualism, and the acquisition of additional languages, see the work of Peter Siemund, most recently Bonnet and Siemund (2018).

Exercises and discussion points

1 Acquisition of phonology

The following table shows a selection of words from the Mali baby-talk register. Describe the adaptations from the adult forms that are made in adapting speech for very young children using the following forms as the basis for your analysis. (Note that the $ symbol indicates a syllable boundary.)

Gloss	Baby-talk word	Adult word
uncle	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]
suckle, feed	[illegible]	[illegible]
eat	[illegible]	[illegible]
father	[illegible]	[illegible]
drink	[illegible]	[illegible]
mother	[illegible]	[illegible]
fish	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]
grandmother	[illegible]	[illegible]
grandfather	[illegible]	[illegible]

2 Acquisition of vocabulary

(a) Below are some examples of inappropriate usage of words that have been cited in the literature on child language. Examine each group, and formulate a hypothesis as to what general meaning the child is attaching to each word.

(i) moon to refer to cakes, round marks on windows, writing on windows and in books, round shapes in books, the letter ‘O’

(ii) tick-tock to refer to a gas-meter, a firehose wound on a spool, a bathroom scale with round dial

(iii) fly to refer to specks of dirt, dust, all small insects, the child’s toes, crumbs of bread, a toad

(iv) chocolate to refer to sugar, tarts, grapes, figs, peaches

(v) door to refer to opening a box, opening a tin, taking limbs off a doll

(vi) bird to refer to sparrows (first referent) then to cows, dogs, cats, any animal moving

(vii) kotibaiz (= bars of cot – the first referent) to refer to large toy abacus, toast rack with parallel bars, pictures of building with columns

(b) In the following examples the young child seems to be using verbs in a way that is not typical of adult usage. Can you find a consistent pattern in these examples?

(i) *I come it closer so it won't fall* (= bring it closer)

(ii) *Daddy, go me round* (= make me go round)

(iii) *I'm gonna fall this on her* (= drop this on her)

(iv) *How would you flat it?* (= flatten it)

3 Acquisition of pragmatics

Use the following questions as a basis for class discussion or reflection.

(a) Recall a telephone conversation you have had with a young child and compare it to the conversation you would have had with the same child face to face (or on Skype). Which is easier? Why?

(b) How does your experience of talking with young children on the phone compare with your experience of talking with older children and adults? What could be some of the reasons for these differences?

4 Development of literacy

Use the following questions as a basis for class discussion or reflection.

(a) How did you learn to read?

(b) Who helped you the most?

(c) What strategies worked best for you?

(d) How was learning to read different from learning to write?

(e) What were the literacy goals behind the teaching you received at school?

(f) What areas of literacy are most important for children today?

5 Second-language acquisition and growing up bilingual

Use the following questions as a basis for class discussion or reflection.

- **(a)** In how many languages can you produce an utterance?
- **(b)** How fluent are you in each of these languages?
- **(c)** How did you learn each one?
- **(d)** Who helped you the most?
- **(e)** Do you think you are good at learning languages?
- **(f)** What strengths do you bring to the task of learning a language?
- **(g)** What challenges have you had to face?
- **(h)** What strategies for language learning suit you the best?
- **(i)** If you were raised bilingually, do you feel like this has benefitted you in any way?
- **(j)** If someone who spoke a language like Spanish approached you for advice about raising their child bilingually, what would you say?
- **(k)** Would your answer be different if they spoke Tongan? If so, why?

Research project

Comparing first- and second-language acquisition

The acquisition of a first language presents an unusual situation where tiny toddlers actually learn better than adults – children acquire any language or languages and they do so with astonishing facility. Research and then outline the many difficulties that second-language learners face compared with those learning a first language, and evaluate the impact of others' expectations on the experience of being a first- or second-language learner. Also, consider what skills older learners can bring to the task and the role of motivation in both types of language acquisition.

For more activities, please go to www.cambridge.edu.au/academic/linguistics.

18

Computational linguistics

◆

Simon Musgrave

18.0 Introduction

Stanley Kubrick's 1968 film *2001: A Space Odyssey* has a computer called HAL 9000 as a major character. HAL is portrayed as using a natural language (English) as freely as any of us. The film is based on a story by Arthur C. Clarke, who later acknowledged criticism that the film did not really have any characters other than HAL. If a computer could in fact communicate like this, that would mean that the major problems addressed by computational linguists would have been solved.

Although we have apps now that have come a long way towards allowing us to interact with technology using our customary ways of speaking, we can't communicate freely with computers in natural language. The app Siri, which is available as an iOS application and allows users to interact verbally with their phones, is one such example. We hope that after reading this chapter, you will have a better appreciation of the considerable achievement such apps represent. This chapter will work through what would have to be going on 'under the hood' in HAL or any similar device and introduce you to major problems in computational linguistics.

There are other linguistic problems for which computers have been used with good results, though, and we will look at some of them briefly towards the end of the chapter.

18.1 How does Hal understand what is said?

The first problem is for the system to be able to break the sound stream into recognisable words. Then, the relationships between the words need to be analysed and a semantic value assigned to the sentence created by these words.

The Turing test

The British mathematician and computer pioneer Alan Turing published a paper in 1950 that established the Turing test by proposing a way to decide whether a computer could think (Turing has been the subject of many movies, including *The Imitation Game*). He imagined a situation in which a human judge could engage in two conversations using keyboard and screen: one with another human and one with a machine. If the judge could not reliably identify which conversation involved a human and which a machine, then Turing claimed it would make no sense to say that the machine was not *thinking*. Turing described this test as though thinking was the only process at issue, but of course linguistic competence is also crucial. The article seems to assume that the language problem is secondary, or perhaps is not even a problem at all. However, 60 years after the publication of the paper, the language problems are not solved. The Loebner Prize has been offered annually since 1991 for a program that can fool human judges, but, up to this point, prizes have only been awarded for the most convincing program in the competition. The 2013 winner, Mitsuku, is capable of carrying on extended interactions, but it is not too hard to see that some chunks of language are produced automatically. We encourage you to go online and try talking to Mitsuku or another chatterbot (or chatbot) and find some of these routines.

Speech recognition

Speech recognition has been the focus of a great deal of research, and the results have been quite impressive, although failures are rather notorious, too (and provide an excellent basis for comedy). But, the problems are very difficult. The acoustic information that is the basis for identifying speech sounds is complex and confusing in various ways, as can be seen when looking at spectrograms (see Chapter 7). Broad classes of sounds can be differentiated reasonably easily: vowels have frequency bands with strong resonance; sonorant consonants also have such bands but with less energy in them; and obstruent consonants (stops and fricatives) lack resonant bands. But, moving beyond this is difficult for various reasons.

- Vowel quality depends on the relationship between resonant frequencies (mainly for the first two formants, as they are known).
- The absolute pitch of these formants varies between speakers (and even between vowel tokens for a single speaker).
- Information about the place of articulation of consonants is available in the changes to vowel formants close to the transition to an adjacent consonant and is not always clear.

We deal with these problems by using a good deal of top-down processing – that is, we can make predictions about what words we expect to hear in many contexts, and these predictions guide us in moving from tentative identification of individual sounds to identifying language chunks. Can computers do the same?

The answer is yes, but in a rather complex way. First, we have to understand that the matching processes deployed here are probabilistic (this is a theme that will continue to develop as we progress). Second, the most successful speech recognition systems use a two-stage process: they first make a probabilistic identification of individual segments, then they assign probabilities to possible word identifications. The first process looks at sequences of sounds and assesses the probability of a certain sound occurring in a sequence. These probabilities are estimated based on a large sample of the language where each segment is known and the probability of different sequences can be estimated accurately. For example, analysing a sample of English in this way will show that the probability of the sequence [s d] at the start of an utterance is zero (the sequence can’t occur at the start of a word). This information is used as new input is processed, with tentative identification assigned to as many segments as possible; the probability of these assignments as part of sequences of segments is then estimated and the most probable identification is adopted (although this may have to be revised as the next segment is assessed, and so on).

The underlying approach here is based on an idea from information theory: the ‘noisy channel’. Any channel of communication contributes noise in transmitting a message and therefore degrades the message. But, if we know something about the properties of the channel (what sort of noise it adds), we can make inferences about the original message. In the case of identifying speech segments, the estimates of the probabilities of various sequences provide information about the channel and these form the basis of inferences about the input message based on the known output message. This concept will turn up several times in this chapter.

The second stage of speech recognition is to recognise words on the basis of a string of (tentatively) identified sound segments. In this case, however, the information against which inferences are tested is much more stable: it is a list of known words (a kind of dictionary, if you like). This word list usually consists of phonemic representations of items. Allophonic variation can be represented by rules in the recognition process, while individual variation has to be accommodated by the statistical parameters of the matching algorithm. In other words, the process the computer follows to work out the correct identity of the word (the matching algorithm) relies on identifying words from lists, but also has to manage the fact that the pronunciation is variable. For this reason, a statistical approach is useful; it allows the computer to make the best approximate matches it can rather than requiring exact matches.

The search space that has to be examined in word recognition is very large, but the problem that causes most errors in such systems is unknown words. Items that are not part of the stored word list tend to cause errors in aligning segments with words elsewhere in the input. Typically, one unknown item will cause between 1.5 and 2 errors in word recognition, so constraining the input is a very important strategy in improving performance in this task (as in many others that we will discuss!). Reported error rates show the impact of this factor: error rates as low as 1 per cent are reported for recognising digits, but word recognition in more spontaneous speech – in this case, a corpus of recordings from switchboard exchanges – has error rates of more than 30 per cent.

Speech recognition in use

Speech-recognition technology is now a part of our everyday lives. Many of us have devices equipped with this technology in our pockets – apps such as Siri and S Voice. Software packages also exist that allow one to control a personal computer with voice commands and to dictate text. Some large corporations require the use of voice-activated software to navigate phone menus. Do you use any of these tools? If so, how well do they work? Speech-recognition techniques are also being used for computer security as an alternative to biometrics such as fingerprint or iris ID. (However, you should certainly take it with a grain of salt when you see a character in a forensic laboratory on television say, 'We identified a suspect from this voiceprint'.)

You might be wondering about the role of phonological and morphological analysis in these processes – we will return to them when looking at the other side of language processing: generating language.

Syntactic and semantic analysis

We have reached the point where an incoming acoustic signal has been turned into a string of words. The next step in speech processing is to look at the syntax of the string, an activity known as **parsing**. As you have seen in a number of earlier chapters, the meaning of constituents larger than individual words is composed in parallel with syntactic composition (in almost all cases); therefore, we have to assign a syntactic analysis to the string in order to compute its meaning. We will consider how this can be done by using rules of the type introduced in Chapter 9; it is also possible to carry out syntactic analysis on a probabilistic basis.

Before parsing can begin, words first have to be assigned to word classes (as described in Chapter 5). This is commonly referred to as **POS (part of speech) tagging**. As we saw in Chapter 9, phrase structure rules have word classes as terminal nodes; therefore, this information is needed before any analysis using such rules can begin. Many words occur in sets of homophones that belong to different word classes – just think of all the English words that can be used as either nouns or verbs – and these ambiguities have to be resolved as far as possible before parsing. For English, if we have a good-sized body of data available to establish relative frequencies, just assigning a word to the word class in which it most frequently appears will get us the correct result about 90 per cent of the time. But, with some additional context taken into account, error rates can be brought lower. It is not hard to see how this will work. For example, if the preceding word is *some*, then the word in question is very unlikely to be a verb; if the following word is a lexical verb marked for tense, then the word in question is very unlikely to be an auxiliary verb.

Rule-based parsing means trying to find a structure for the input string that matches the available rules and word class labels. We will look at a simple example in rather excruciating detail in order to help you see that even a ‘simple’ example is not so simple, and also to help you understand the number of steps involved in natural language processing of this kind. We will look at the input string:

The man ate a chocolate cake

And we will assume the following phrase structure rules and lexical insertion rules:

S → NP VP
NP→ Det (N) N
VP → V (NP)
N: [man, chocolate, cake]
DET: [the, a]
V: [ate]

In the above rules, S = a clause, N = noun, V = verb, Det = determiner, NP = noun phrase and VP = verb phrase. Recall that POS labelling has already been completed, so we know the part of speech for each word. Applying a parsing algorithm to the input string *The man ate a chocolate cake* might result in the following steps.

1 Take first word: the_Det.

2 Check whether word on its own can be RHS (right-hand side) of any rule: NO.

3 Check whether word is first element of the RHS of any rule: YES.

4 Assume NP.

5 Take next word: man_N.

6 Check whether word can be next element in NP: YES (but note that in the rule above, this is the optional N in the brackets).

7 Continue to assume NP.

8 Take next word: ate_V.

9 Check whether word can be next element in NP: NO.

10 Parse fails, start again.

11 Take first word : the_Det.

12 Check whether word on its own can be RHS of any rule: NO.

13 Check whether word is first element of the RHS of any rule:

YES.

14 Assume NP.

15 Take next word: man_N.

16 Ignore optional element.

17 Check whether word can be next element in NP: YES.

18 Check whether rule requires another element: NO.

19 Assign label NP to substring The man.

20 Check whether NP is first element of the RHS of any rule: YES.

21 Look for VP.

22 Take next word: ate_V.

23 Check whether word can be only element in VP: YES.

24 Assign label VP to substring ate.

25 Check whether expansion of root node has been completed: YES.

26 Check whether all of input string has been used: NO.

27 Parse fails, start again.

28 … … … …

Even at this level of detail, this is very tedious (and, of course, an actual implementation of an algorithm [routine] like this would have to specify many of the steps in even more detail), as well as being repetitive and inefficient. Computational linguists have devoted a good deal of ingenuity to making parsing more efficient. One strategy that has been successful is already implied by step 16 above: in our example, we assumed that the program 'knew' not to try a possibility that had already failed. But, of course, a computer has to be told explicitly to store this information and to use it at the appropriate time. A program known as a chart parser does this systematically. Dead ends that have been tried are recorded in a data structure (a chart) along with the point at which another choice was possible. When the parser comes to that point again, it has the information about previous failures available and can try another analysis.

Steps such as 3, 13 and 20 above suggest another strategy that might increase the efficiency of a parser. Storing information about the possible first elements in various constituents will reduce the number of possibilities to be examined. In our simple example, not much would be gained by storing the information that encountering a determiner must mean the start of a noun phrase, and encountering a verb must mean the start of a verb phrase, although even that information would provide some gains. If we have more complex rules and more complex input, the advantages of this sort of approach are considerable. Keeping a list of the so-called left corners of constituents can make a parser considerably more efficient.

Algorithms

In the early 1970s, Professor Harold Stone defined an algorithm as ‘a set of rules that precisely defines a sequence of operations’. This covers all computer programs – Stone extended his definition to include ‘any sequence of instructions that can be obeyed by a robot’. The following example is an informal (that is, ordinary-language) algorithm to find the largest number in an unsorted list.

1 Assume the first item is the largest number and note it.

2 Go to the next item.

3 If that item is larger than the noted number, cross out the noted number and note the new larger number.

4 If there are more items, return to step 2; otherwise, the last noted number is the largest in the list.

Some features here are immediately striking. First, the level of explicit detail is greater than we (as non-robots) need (try reformulating these instructions for ordinary use). Second, we can see that structures common in computer programs (if/then tests, loops) are necessary in even simple algorithms. And third, it is very hard to specify even a simple algorithm like this one without assuming more basic algorithms – here we assume one that tests which number in a given pair is larger. (What might that algorithm look like?)

The example above is sufficiently simple that even our basic rules only allow one parse for the string. But, this is rarely the case: any rules that are general enough to give good coverage of the possibilities of a language will allow multiple parses in many cases. Building a robust parsing system means trying to limit the possibility of multiple parses while not making the rules so narrow that they will fail to parse acceptable chunks of the language. For example, we can slightly modify our simple rules to allow that the determiner in a noun phrase is optional, accommodating examples such as *Wise men like sheep* or *Language is useful*. If we also allow that the element preceding the head noun can be a phrase (for example, *a [white chocolate] cake*), then we would have the rule:

NP → (DET) (NP) N

But, now when we try to parse even our original simple example, we will get two results. In one, the determiner will be part of the larger noun phrase; in the other, it will be part of the optional noun phrase. In order to prevent this from happening, we have to be able to specify when it is possible to omit the determiner from a noun phrase and how this can be included in a rule system. (We encourage you to think about these two questions – they are not easy!) Of course, there are cases where we want our parsing system to produce more than one result, such as where there are structural ambiguities of the sort introduced in Chapter 9.

Knowing the syntactic structure of an utterance means that its meaning (or at least part of it) can be computed. One step may be necessary before doing this, and that is resolving lexical ambiguities. Even when we know the part of speech to which a word belongs, more than one sense may be associated with that part of speech and knowing which sense is relevant is necessary in establishing meaning. You have seen in Chapter 6 various ways in which the meaning of linguistic expressions can be captured. We won't worry too much here about how meaning might be represented, but it is important to understand that the aim is to achieve a representation that allows a computer to make valid inferences on the basis of the language input. Semanticists have devised systems of representation that allow some aspects of the context of use to be included in making inferences, but including the general knowledge and commonsense understanding of the world that we bring to interpreting linguistic expressions is usually considered to be part of a more general problem in research on artificial intelligence. This is the problem (sometimes called the frame problem) of how much knowledge about the world needs to be encoded in a computer before it can respond to situations in a way that we might see as intelligent.

18.2 How does HAL say anything?

For our present purposes, we will assume that HAL (or whatever machine we are examining) has made some inference on the basis of the linguistic input that has been processed, and that a response can be based on that inference. The response will start out as a semantic representation; this will be the basis for the selection of items from the lexicon that will be combined into a syntax structure using the same rules used in parsing input. The output of this process will be a string of morphemes. Phonological rules will then have to be applied to this string to create a string that can be turned into sound. This is similar to the process described for human speech generation in Chapter 16.

The most common approach to implementing phonological rules in computational linguistics is to use finite state networks. These consist of nodes, representing states, and arcs between them. Transition from one state to another depends on meeting some condition. In the case of phonological rules, the condition consists of a correspondence between a phone (speech sound) in the input and a phone in the output. Permitted sequences of correspondences are represented as paths through the network; there are no paths for prohibited sequences. We can exemplify this with the (partial) network in Figure 18.1 showing the allomorphy associated with the plural morpheme in English, as described in section 8.5 of Chapter 8. At the left-hand side, there is an entry point that can be any sound forming the penultimate phone in a base noun. There are paths leading from this state to three other states; one arc is traversed if a voiced sound in the input corresponds to the same voiced sound in the output; another arc is traversed if a voiceless sound in the input corresponds to the same voiceless sound in the output; and the final arc can be traversed if a sibilant sound in the input corresponds to a sibilant in the output. Then, there is a path to the word boundary from each of those three states. In the first case, reaching the word boundary requires having an [s] in the input (we’re assuming that as the basic plural morpheme) corresponding to a [z] in the output; in the second case, the input [s] corresponds to the output [s]; and in the final case nothing (shown as a zero) in the input corresponds to [ə] in the output and the arc leads to the state which was otherwise reached by having a voiced segment. From there, the end of the word can be reached with an [s]/[z] correspondence. So, the network allows only the sequences of correspondences shown here:

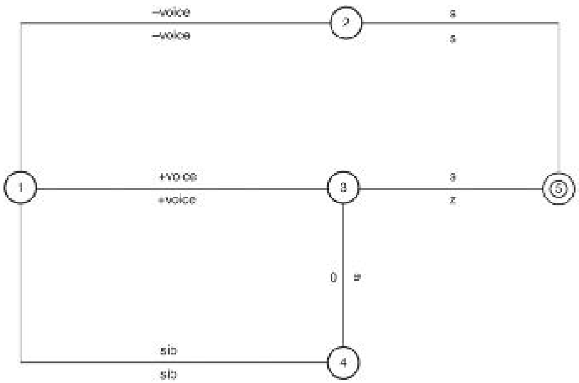


Figure 18.1 A finite state network representing part of the allomorphy of the English plural

[+voice][z], [–voice][s], [+sibilant][ə][z]

You can check for yourself that there are no other possible paths.

The output of this stage of processing will be a phonetic representation of an utterance that can be the basis for synthesising speech sounds. There are two ways of approaching this last problem: one is to really synthesise speech sounds from scratch; the other is to combine pieces of sound recorded from real people speaking to create the new utterance. For many of the same reasons that speech recognition is so difficult, the first approach – synthesising speech sounds from scratch – is not very satisfactory: the **speech synthesis** box below gives some more information, and directs you to some websites and software demonstrations that will let you hear the sort of results that have been achieved. Concatenating recorded chunks produces more natural-sounding speech, but brings its own problems. The most obvious of these is deciding what size of chunks to use. The answer to this question depends very much on what we are trying to do.

We have all had the experience of interacting with automated telephone-answering systems; what we hear from such systems is pre-recorded and it is usually easy to tell when smaller chunks are being used. A system that allows you to pay bills on the telephone will typically have all instructions as complete prefabricated units up to the point where you are asked to confirm a number that you have entered. At this point, you are likely to hear something like, 'Did you say [string of digits]', where the first three words have been recorded as a unit and sound natural, but the digits have each been recorded individually and sound quite different. The main way in which the string of digits sounds odd is that the intonation contour of a word is presented in isolation, whereas if we read a string of digits, we impose some grouping on the digits and this changes the intonation pattern.

This makes us aware of a major problem in making any speech synthesised from small units sound natural. Assigning intonation to an utterance so that it sounds natural is another complex problem; in most cases the best that can be achieved is a neutral intonation that doesn't sound obviously wrong.

Speech synthesis

The human vocal tract has complex acoustic properties and this makes it hard to synthesise speech sounds that sound natural. The demonstration that you can download from the Centre for Speech Technology KTH Stockholm produces vowel-like sounds; you can manipulate the formant frequencies and other acoustic properties of the sound, but importantly you can choose between basing the sound on the electronic source or on sampled human speech. Neither sounds very good in this decontextualised environment, but we think the human sample base is a lot better. The websites listed in the online resources for this chapter allow you to hear speech synthesised from a text input of your choosing, with various possible voices and accents.

We have now worked through the numerous problems that would have to be solved in order for a computer to achieve the behaviour of which HAL is depicted as being capable. We hope that, from this, you have gained an appreciation of just how difficult the problems are, and why interaction between humans and computers using natural language is not yet very successful. There are other linguistic problems for which computers have been used with good results, though, and we will now look at some of them briefly.

18.3 Machine translation

A great deal of research has been devoted to **machine translation**, using computers to translate texts from one language into another. Before about 1990, work in this area employed approaches based on rules, and much of the previous discussion about how HAL might work is relevant. We can think of an input being processed as described in the earlier part of this chapter, but, then, instead of using the result as the basis for inference and generation of a response, the analysed input is used as the basis for generating an equivalent text in a different language. Systems using this sort of approach vary in the level at which the transfer between languages takes place. It may be done on the basis of a transfer between language-specific semantic representations, or on the basis of a more general (possibly universal) semantic representation.

You can get a sense of the issues involved here by thinking again about the example of English determiners that was mentioned earlier. Many languages don’t have the same requirement as English for definiteness to be specified for all noun phrases. How, then, might transfer of meaning happen between English and such a language? It is not too difficult to see how going from English to the other language might work – we can discard irrelevant information – but going in the other direction is more complex: if the information is not represented in the source language, where should we look for it in order to represent it (as we must) in English? As with other areas already discussed, limiting the kinds of input that have to be handled leads to better results for machine translation. One of the best known and most successful rule-based translation systems is ‘METEO’, which was developed by Environment Canada to translate weather forecasts between English and French. This task can be limited to a small vocabulary and range of sentence structures, and results were thus very good, although, even then, forecasts were checked by human translators before publication.

In recent decades, work on machine translation has concentrated on probabilistic approaches. Most of the online translation engines (such as Google Translate or Babel Fish) are based on such approaches. This strategy involves large collections of parallel sentences in a pair of languages. The sentences are analysed by a computer to establish the probability that a word or a sequence of words in the source language corresponds to a word (or a sequence) in the target language. This process is known as training and the collection of data on which it is based is training data. When provided with new source language input, the translation engine looks for the sequence of words in the target language that has the maximum probability of corresponding, based on the parameters established in the training phase. The success of this approach depends very much on the amount and quality of the training data; the timing of the start of serious work along these lines is due partly to the fact that computer memory has become much cheaper in recent years. This approach to translation is far from perfect, but the fact that this approach to translation works even as well as it does raises questions for how we think about our knowledge of language.

Round-trip translation

If you experiment a little with a service like Google Translate, it isn't too hard to get an idea of whether or not a large body of training data exists for a pair of languages. One way of testing this, even for languages that you do not know, is to attempt round-trip translation. To do this, input some text in English and translate it to any other language, then use the output of that translation as the input for a translation back into English. Ideally, the final result should be identical to the original text, but this is rarely the case! Probabilistic machine translation does not represent any aspect of the structure of language in its processes. This means that errors of word order are not uncommon. For example, round-trip translation via Hindi produced the following output based on the source text.

Input text

> Police have vowed to hunt down those responsible for yesterday's violent protest in Sydney's CBD in which six officers were injured.

Output text

> Police vowed to hunt down those who were six officers injured CBD yesterday's violent protests in Sydney is responsible for.

18.4 Corpus linguistics

Another area of research that we mention here is the field known as **corpus linguistics**. As discussed in Chapter 3, the word *corpus* can be used to describe any collection of data, but in the phrase *corpus linguistics* it is normally understood to mean a large collection of data in a form that can be processed by computers. Corpus linguistics, in this sense, has existed for at least 50 years, but as with probabilistic machine translation, access to cheap memory (and more powerful computers) has meant that the field has expanded rapidly in recent times. The extent of this expansion can be seen in the size of the collections of data that are being used. One of the first important research corpora (corpora is the plural form of corpus) was the Brown Corpus compiled in the 1960s and consisting of approximately one million words. By the 1990s, the British National Corpus had been assembled, consisting of 100 million words. Today, we have online access to corpora such as Global Web-Based English (1.9 billion words) and the Google Books corpus (34 billion words of British English).

Corpus linguists investigate many different questions about language, but the basis for most such work is the idea that looking at a sufficiently large body of data will allow us to detect patterns that we might not otherwise observe, and that finding such patterns depends on the capability of computers to search large bodies of data quickly and exhaustively. To see how this can work, let's consider a simple question involving the word *different*. Some people say *different from*, some say *different to*. There is also the possibility of *different than*. There are disparities in usage (at least in relative proportions) between varieties of English in, for example, Great Britain and the United States. But, the question we will look at is: What forms are used by speakers for whom English is not their first language? We are fortunate that there is a freely available corpus of data relevant for examining this question, the Vienna-Oxford International Corpus of English (VOICE). Figure 18.2 is a screenshot from the corpus analysis software AntConc, showing the most basic kind of analysis of corpus data: word frequencies. We see from this that the word *different* is the 107th most common word in this corpus, with 1331 occurrences in 1 024 974 words of text (which can be normalised to 1298 occurrences per million words). We could compare this with data from the British National Corpus (BNC): in that collection, *different* is the 100th most common word and occurs 484 times per million words. We can already see that there is some variation in usage between these two corpora; although the rank frequency of the word is similar between the two, the rate of occurrence is quite different.

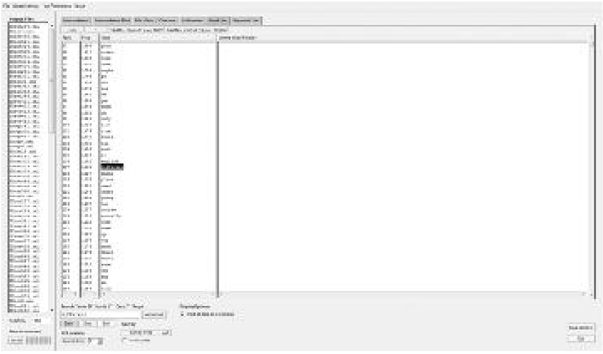


Figure 18.2 Word frequencies

Figure 18.3 shows a screenshot from AntConc's Key Word in Context tool. This tool lets us see each occurrence of a word with the text that surrounds it. In this case, we have sorted the list (called a **concordance**) by the following word so that all the *different from* examples are together.

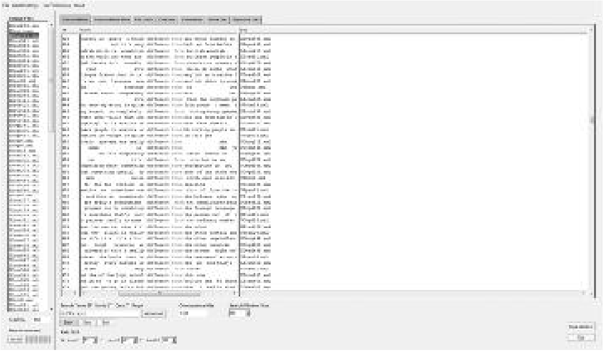


Figure 18.3 Word occurrences

Figure 18.4 is a screenshot of the results from another AntConc tool; this one looks at collocates, or words that occur together. From these results, we can see that there are 50 occurrences of *different from*, seventeen occurrences of *different than* and eight occurrences of *different to*. Again, we can compare these results with data from the British National Corpus. A sample of 1000 occurrences of the word *different* in that source gives 76 *different from*, only one *different than* and ten *different to*.

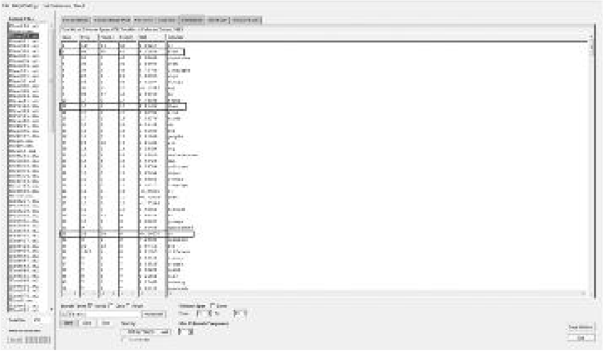
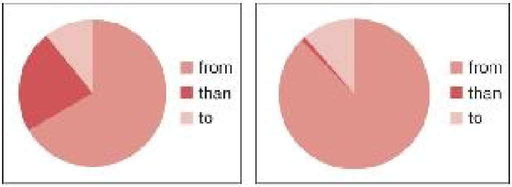


Figure 18.4 Word collocates

Figure 18.5 shows this comparison in chart form; it is clear that there are differences in usage in these corpora. It is also, we hope, clear that uncovering these sorts of differences without the assistance of computational tools would be difficult (and very time-consuming).



Different in VOICE *Different* in BNC

Figure 18.5 Comparison of *different* in VOICE and BNC

Further analysis of examples in context is required to fully understand the nuances in the patterns identified in this way, and one way to make the results of such analysis accessible is to show the relationships between collocates as a network diagram. Figures 18.6 and 18.7 show the contrasting **collocations** of *heaps* and *lots* in New Zealand English. As Andrea Calude notes in her research, these diagrams help us to see that although the two networks have many features in common, there are many more verbs associated with *heaps* than with *lots*.

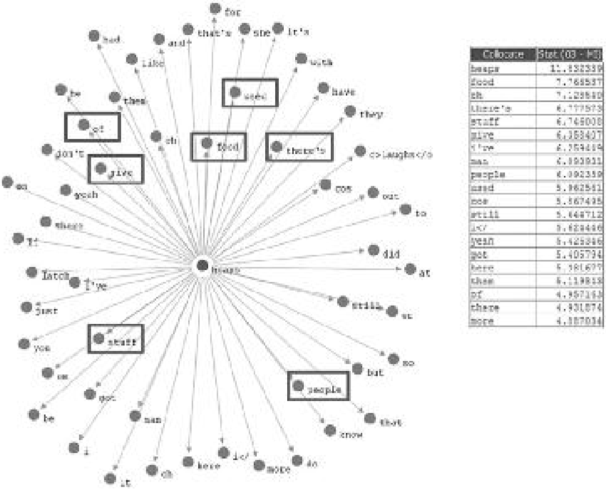


Figure 18.6 Collocations of *heaps* in New Zealand English

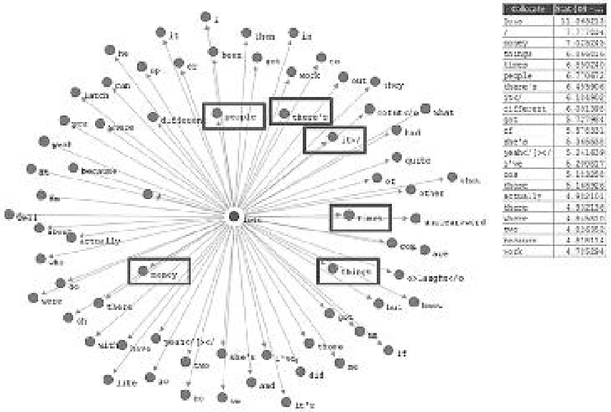


Figure 18.7 Collocations of *lots* in New Zealand English

These examples should give you some idea of the kinds of questions that can be investigated using the techniques of corpus linguistics, but the possibilities are much more exciting than this. These techniques can be used to look at what linguistic features characterise particular types of text (see following section), to test attributions of authorship, to trace chronological patterns of usage and address many other problems. Another very important use of corpus data is as a basis for describing language in relation to actual usage. All modern dictionaries are based on corpus data, and there are now several excellent grammars of languages that have been developed in this way.

Corpus linguistics and discourse analysis

In addition to analysing and exploring word frequency and collocations, computational tools for corpus linguistics are increasingly used in projects focused on discourse analysis, which explores patterns in language above the level of the clause or sentence, as well as a range of questions about how language is used for particular social purposes (see also Chapters 13 and 15). Here we will briefly introduce some of the techniques used in this research.

An important question in analysing the purpose and impact of a text may be whether certain words characterise the text, or whether certain words are absent from the text. **Keyword analysis** uses statistical methods to compare the frequency of occurrence of words in one text or collection of texts (a target corpus) and another collection of texts (a reference corpus), and gives a measure of how unexpected the frequency of a word in the target corpus is, given its frequency in the reference corpus. For example, comparing material from a conservative news source with material from a range of other news sources might show that the word *family* occurred more often than we would expect in the target corpus, but that the word *partner* occurred less often than expected. Keyword analysis can tell us something about the nature of a text on its own, but it is often used to identify words whose function in the text should be studied in more detail.

Various methods can be used to extend this type of analysis to look at semantic fields rather than individual words. One possibility is to use automated tagging to assign words to semantic fields; the resulting tags can then be the input to a keyword analysis. Rather than looking at unexpected frequencies for words, this procedure will look for unexpected frequencies of the tags that have been assigned. In recent years, researchers have also used machine-learning methods to look for semantic patterns in texts. The method known as topic modelling (more formally, Latent Dirichlet Allocation) analyses a group of texts and looks at the patterns of co-occurrence of words in texts. The output of the analysis is groups of words ('topics') that tend to occur together, with measures of the strength of relation between (a) topics and texts, and (b) topics and words. The word-embedding method (see box below, 'Big Brother is coming!') is also very powerful for identifying semantic relations, as it produces a model of a text sample in which the distances between words represent semantic relations, and those distances can be given a precise numerical value.

We have already mentioned collocation analysis as a basic technique in corpus linguistics. Collocates are words that tend to co-occur close to each other, but they are not necessarily fixed phrases. Collocation analysis only looks at the relations between two words at a time; to find fixed phrases that extend over more than two words, we can use lexical cluster analysis (lexical clusters are often referred to as *n*-grams, where *n* is the number of words being looked for). Such analysis is interesting because it shows the extent to which spontaneous use of language depends on pre-assembled chunks (such as *the best way to* … or *What do you think about* …), and provides insights into how discourse is structured. We are typically not aware of these chunks and instances of them are often incomplete.

It is possible to use several of these techniques together; such approaches have been found to be especially relevant in classifying texts. For example, multi-dimensional analysis is a technique that was first developed by Douglas Biber. This method combines information about a range of features of texts (for example, mean word length, frequency of occurrence of different word classes and use of different ways of combining clauses) to produce measures that reliably identify texts as belonging to particular registers and genres. Biber's analyses already use sophisticated statistical methods, but it is possible to go even further in that direction and use regression analyses to identify interactions between linguistic and contextual features. An example of such a method is the variable rule procedure used by some sociolinguists.

Big Brother is coming!

A fascinating new area of research at the intersection of language and computers has emerged in the last few years, the study of *word embeddings* – also known as *vector space models*. In this research, machine learning methods are applied to large bodies of text data – a seminal paper in the field by Mikolov and colleagues describes a model built on 1.6 billion words of text to produce multi-dimensional (think of 100 or 200 dimensions) models that locates every word in the text in relation to every other word. Such models are mathematically precise and turn out to have very interesting properties as a result. For example, excellent performance on analogy problems is achieved with simple arithmetic operations: subtracting the series of numbers that represents *king* from the series of numbers that represents *queen*, and adding the result to the series of numbers that represents *prince*, will get you close to the position of *princess* in the model. Such methods can then be used, for example, to provide important results on how language and language use are gendered; the American scholar Ben Schmidt has a fascinating (and scary!) report of his work on data from the *Rate My Professor* website along these lines. Perhaps even scarier is very recent work by Fernando Pereira and his colleagues, which shows that computers can be trained to take meanings as represented by the mathematical objects in a vector space model and associate them with magnetic resonance images of brain activation. Does this mean a computer can already tell what you are thinking?

18.5 Techniques from genetics

A final example of the use of computers in linguistic research is the use of **clustering algorithms** developed for research on genetics to analyse language data. The amount of genetic information in the DNA of living organisms is enormous, but comparing the similarities and differences between this information for different populations is a crucial process in genetics research. Consequently, geneticists have developed powerful computational tools that carry out these comparisons and then use them as the basis for generating models of the relationships between populations and organisms. In recent years, these tools have also been applied to language data with two main aims. First, linguists have been keen to apply such procedures to data from language families (such as Indo-European) that have been extensively studied in order to see whether the computational results agree with the picture built up by traditional comparative linguistics. The good news for linguists is that, by and large, the results from the two approaches do correspond well. This provides the basis for the second way in which these tools have been applied in linguistics. They have also been used to approach problems that, for various reasons, had seemed resistant to the traditional approaches. We will present an example of such research in a moment, but first we should mention some issues that must be borne in mind when considering results of this type.

It is important to understand that the type of work that we are discussing here still depends on human judgement in the preparation of the data to be analysed. The vocabulary of DNA has only four characters (G, C, A, T), and they can only occur in two combinations (GC, AT); therefore, it is completely feasible for all comparisons to be accomplished by a machine. A typical human language has around 30 characters (phonemes) that can occur in a large (but not unrestricted) number of combinations. To date, the data that form the input to the clustering algorithms used on similarities and differences between languages have to be produced by human judgement. These judgements may be open to dispute or even to error. The process itself is time-consuming, and this is also a limitation on the amount of data that may be available for the analysis. This point also leads to a more general issue about the quality of data that is used in such analyses. There have been several instances in recent years where research has been published in prominent journals claiming important results from the computational analysis of historical language data, and the level of linguistic expertise of the researchers and the quality of the data on which the analysis is based have been open to question. (Thomason's work is a recent critique of such an example.) An excellent introduction to the type of research discussed here is the 2005 book by April and Robert McMahon; it was written by an eminent linguist and an eminent geneticist (who happen to be married to each other), but such balanced collaboration is not always seen. With these caveats in mind, we will now briefly describe one fascinating piece of research in this tradition.

Given the devastating impact of European invasion on the Aboriginal population of Tasmania, we know very little about the languages spoken on the island before the arrival of the British. By the 1830s, the languages were no longer used in everyday communication and the only data available now on these languages are word lists and vocabularies; estimates of the number of languages have varied between five and sixteen, and there has been no agreement about whether these were members of a single family or of several groups. The Australian linguist Claire Bowern applied the techniques we have introduced here to this word-list data in a study published in 2012 and presented results that clarified the situation in several important areas. First, Bowern is able to make an assessment of the extent of mixing of different languages in some of the records and this leads, in turn, to the possibility of identifying the primary language recorded in lists that were previously seen as unclassifiable. Second, the algorithms employed by Bowern suggest that the most probable result is that there are twelve distinct languages represented in the extant data, and that these form five related clusters that correspond well to geography and to the archaeological record. However, only two of these clusters can be related to each other at a higher level and the analysis gives no grounds for viewing the Tasmanian languages as a single family. In this case, the employment of powerful computational methods has produced answers to long-standing questions that are more reliable than those given by earlier comparative linguists. Such techniques are being applied more and more frequently in linguistics, and will certainly produce more results as interesting and important as those of Bowern.

18.6 Conclusion

In this chapter, we have tried to introduce you to some of the ways in which the field of linguistics has intersected with the use of computers. We have only touched on some of the main areas of research; whole areas have not been mentioned. Nevertheless, we hope that we have given enough information to arouse your interest, and that you will feel inclined to look at some of the online resources referred to, or to experiment with software tools like AntConc. We also hope that some of what you have learnt in this chapter prompts you to revisit some basic questions about language. Computers require explicit instructions in order to carry out tasks; thus, what does it mean that we struggle to provide instructions about how language is used that are sufficiently explicit to enable a computer to talk?

Although we probably do not process language probabilistically, what do we learn about our capacities from the fact that some language-related tasks can be handled by computers on a purely statistical basis?

Further reading

Many useful resources relating to this chapter are available in the online materials for this book. Clarke discusses HAL in his 1972 book. Turing (1950) has been reprinted often; a good place to find this article and discussion on it is in Hofstadter and Dennett (1982), where there is an important response to Turing from John Searle as well as editorial commentary. Section IV of Dennett (2013) provides a very clear introduction to the basics of how the power of computers is based on algorithms. This text includes practical exercises that use the software tool RodRego (see Exercise 3 below). The switchboard exchange data is from Jurafsky and Martin (2008), which includes an excellent overview chapter (Chapter 24) on machine translation; it is quite technical but also very clearly written. If you prefer a more philosophical approach, Douglas Hofstadter's 1997 book *Le Ton Beau de Marot: In Praise of the Music of Language* is a long meditation on the problems of translation that includes several computer-generated translations of a (short) French poem. McEnery and Hardie (2011) is an excellent introduction to current practice in corpus linguistics. For multi-dimensional analysis, see Biber (1988) and for word embeddings, see Mikolov et al. (2013). As we noted, McMahon and McMahon (2005) provides an excellent overview of applying computational methods from genetics to languages. Thomason (2013) outlines a number of practical problems with this approach, and you can read more about the application of these techniques to the Tasmanian languages in Bowern (2012).

It is easy to access large bodies of language data today, as there are many excellent resources available online. We offer just a few suggestions here, but it is not hard to find many other useful web resources. A good place to start is the Google Books Ngram Viewer: https://books.google.com/ngrams. This tool gives you a visual representation of the frequency with which words occur over time in the data scanned by Google Books.

The website https://www.english-corpora.org provides an online interface to several large corpora, including the British National Corpus and the Corpus of Contemporary American English. The search interface is sophisticated and allows you to see keyword-in-context results, collocates and clusters (n-grams). WebCorp Live states on its website (http://www.webcorp.org.uk/live) that it is ‘a suite of tools which allows access to the World Wide Web as a corpus – a large collection of texts from which facts about the language can be extracted’. To explore the range of methods available in corpus linguistics see Baker and Egbert (2016). The examples of semantic network analysis are from Calude (2017). The work on modelling brain activation is Pereira et al. (2018).

Several of the figures in this chapter are screenshots from the software package AntConc. (http://www.laurenceanthony.net/antconc_index.html). This is a free package that is powerful enough to do most things corpus linguists want to do, but it is quite simple to learn to use.

Exercises and discussion points

1 Chatbots

The first well-known chatbot (although it wasn’t called that at the time) was ELIZA, which claims to be a psychotherapist. There is an implementation at the website manifestation.com. Mitsuku was mentioned earlier in the chapter; other chatbots include:

- Elbot (2008 Loebner Prize winner)
- Ultra Hal Assistant (2007 Loebner Prize winner)
- Jabberwacky.

Look up each of these chatbots using the website addresses provided in the online resources for this text. Engage in a conversation with the bots. Are there any noticeable differences between the bots’ language and human language? What do these differences tell us about the way that chatbots and machine generated language work?

2 Speech recognition

Using your own device, or sharing with someone who has one, try to fool the speech recognition app. How much does successful performance depend on factors such as speed of speech, background noise and accent?

3 Algorithms

RodRego is an implementation of a very simple computing machine called a register machine. You can use it to try the problems Dennett includes in his 2013 text and, generally, to get a sense of how algorithms work in practice. It is only fair to warn you that some people find this exercise mind-numbingly boring and give up very quickly, while others find it fascinating and turn away from it after what seems a short time to find that they have missed several meals!

4 Speech synthesis

AT&T Labs has a small demonstration online that handles English and Spanish with various voices. Acapela Group offers speech synthesis for many more languages and with many voice possibilities. (It is fun to input English text and have a completely different language as the basis for the sounds.)

Try different possibilities from the two websites provided in the online resources. Which ones do you find most convincing? (Make sure to try the Australian English voices at Acapela.) What makes some examples of synthesised speech sound better than other examples?

5 Round-trip translation

Experiment with one or more of the online translation services such as Google Translate or Babel Fish. Start by trying translation between two languages that you know. How good are the results? What causes problems? Then, move on to round-trip translation and see what sorts of results you get. For an intermediary language that you don't know, can you guess anything about the properties of the language (such as word order) on the basis of the final English result?

Research project

Doing corpus linguistics

In this activity, we will guide you to an extension of the analysis of the word *different* that was presented earlier in the chapter. We will work with data from the British National Corpus (BNC) and the Corpus of Contemporary American English (COCA) available from the www.english-corpora.org website. You will need to undertake several searches, but the general outline will be the same for each.

> Choose **Display** as KWIC.
>
> Enter **Word** *different*.
>
> Ignore **Sections** for now.
>
> In **Display/Sort, Display** = ‘Alphabetical’ and click in the box one space to the right of the centre dark blue box for **Sort**.
>
> For **#Hits**, select KWIC = 1000

These settings will give you a sample of 1000 occurrences of the word *different* sorted alphabetically by the following word, making it easy to count how many occurrences of following *from, than* and *to* there are. You should run this search five times for each corpus and record the relevant numbers for each search in a table. Then start restricting the search by ‘Section’ – that is, search only ‘Spoken’ texts, then only ‘Fiction’ texts, and so on. Use the COCA categories only; the BNC categories allow for more fine-grained categories but there is no way to compare these with COCA. Remember to keep recording the numbers of occurrences of the different patterns as you go. Ideally, you will do each search several times, but even if you only search each sub-section once, you should still get useful results. When you have completed all the searches, look at the results you have recorded in your table and start trying to make sense of them. Below are some questions that you should try to answer.

1 What is the most useful way to compare these numbers? Are the raw numbers easily comparable or would it make them more comprehensible to manipulate them in some way?

2 Are there overall differences between British usage and American usage? If so, what are they?

3 Are there differences in usage between text types? Do some text types look more or less similar between the two varieties, and if so, what might this tell us about how language is used?

4 How similar are the rates of occurrence of the different patterns across the five samples you took from each corpus as a whole? Is there a level of variation in either corpus that might make us worried about the reliability of our results?

References

Abner, N., Cooperrider, K. & Goldin-Meadow, S. 2015. ‘Gesture for linguists: A handy primer’. *Language and Linguistics Compass*, **9**(11), 437–51.

Aboud, F. 1976. ‘Social developmental aspects of language’. *Papers in Linguistics*, **9**, 15–37.

Adams, M. 2009. *Slang: The People’s Poetry*. Oxford: Oxford University Press.

Adams, R.M. 1973. *Proteus, His Lies, His Truth: Discussions of Literary Translation*. New York: W.W. Norton.

Adesope, O.O., Lavin, T., Thompson, T. & Ungerleider, C. 2010. ‘A systematic review and meta-analysis of the cognitive correlates of bilingualism’. *Review of Educational Research*, **80**(2), 207–45.

Ainge Roy, E. 2018. ‘Māori has gone mainstream’: The resurgence of New Zealand’s te reo language’. *The Guardian,* 28 July, https://www.theguardian.com/world/2018/jul/28/google-disney-maori-new-zealand.

Aitchison, J. 1992. *Linguistics: Teach Yourself Books*. London: Hodder and Stoughton.

Aitchison, J. 2003. *Words in the Mind*. Oxford: Blackwell.

Aitchison, J. 2007. *The Articulate Mammal: An Introduction to Psycholinguistics*. London: Routledge.

Aitchison, J. 2013. *Language Change: Progress or Decay?* 4th edn. Cambridge: Cambridge University Press.

Allan, K. 1986. *Linguistic Meaning,* Vols 1 and 2. London: Routledge and Kegan Paul.

Allan, K. 2001. *Natural Language Semantics. Oxford*: Blackwell.

Allan, K. 2010. *The Western Classical Tradition in Linguistics*. 2nd edn. London: Equinox.

Allan, K. & Burridge, K. 1991. *Euphemism and Dysphemism: Language Used as Shield and Weapon*. New York: Oxford University Press.

Allan, K. & Burridge, K. 2006. *Forbidden Words: Taboo and the Censoring of Language*. Cambridge: Cambridge University Press.

Allan, K. & Burridge, K. 2010. 'Swearing and taboo language in Australian English'. In P. Peters, P. Collins & A. Smith (eds), *Comparative Grammatical Studies in Australian and New Zealand English* (pp. 295–347). Amsterdam: John Benjamins.

Allen, W.F. 1987. '*Teenage speech: The social dialects of Melbourne teenagers'*. Unpublished BA Hons thesis. Bundoora: La Trobe University.

Altenberg, B. 1990. 'Speech as linear composition'. In G. Caie, K. Haastrup, A. Jakobsen, J. Nielsen, J. Sevaldsen, H. Specht & A. Zettersten (eds), *Proceedings from the Fourth Nordic Conference for English Studies*, vol. 1 (pp. 133–43). Copenhagen: University of Copenhagen.

Ambridge, B. & Lieven, E.V.M. 2011. *Child Language Acquisition: Contrasting Theoretical Approaches*. Cambridge: Cambridge University Press.

Anderssen, L.G. & Trudgill, P. 1990. *Bad Language*. Harmondsworth: Penguin.

Andrews, M., Squire, C. & Tamboukou, M. (eds) 2008. *Doing Narrative Research*. London: Sage.

Anonymous n.d. 'Maori nation – What Maori nation?' *Wordpress*, http://onenzfoundation.co.nz/wordpress/articles/treaty-of-waitangi/maori-nation-what-maori-nation.

Aristotle 1984. *The Complete Works of Aristotle: The Revised Oxford Translation*. Edited by J. Barnes. Princeton, NJ: Princeton University Press.

Arnold, B.B. 2012. 'Barnold Law – Names'. *Blogspot*, http://barnoldlaw.blogspot.com.au/2012/01/names.html.

Austin, J.L. 1962. *How to Do Things with Words*. London: Oxford University Press.

Australian Institute of Aboriginal and Torres Strait Islander Studies (AIATSIS) & Federation of Aboriginal and Torres Strait Islander Languages (FATSIL) 2005, *National Indigenous Languages Survey Report 2005*. Canberra: Department of Communications, Information Technology and the Arts.

Ayçiçegi-Dinn, A. & Caldwell-Harris, C.L. 2009. 'Emotion-memory effects in bilingual speakers: A levels-of-processing approach'. *Bilingualism: Language and Cognition*, **12**(3), 291–303.

Baker, C. 2000. *The Care and Education of Young Bilinguals*. Clevedon: Multilingual Matters.

Baker, P. & Egbert., J. 2016. *Triangulating Methodological Approaches in Corpus Linguistic Research*. London: Routledge

Baldwin, J. & French, P. 1990. *Forensic Phonetics*. London: Pinter.

Barker, C. 2000. *The Care and Education of Young Bilinguals*. Clevedon: Multilingual Matters.

Barske, T. & Golato, A. 2010. ‘German *so*: Managing sequence and action‘. *Text & Talk: An Interdisciplinary Journal of Language, Discourse & Communication Studies*, **30**(3), 245–66.

Battistella, E. 2006. ‘Girly men and girly girls’. *American Speech*, **81**(1), 100–10.

Bauer, L. 1983. *English Word-formation*. Cambridge: Cambridge University Press.

Bauer, L. 1994. *Watching English Change: An Introduction to the Study of Linguistic Change in the Twentieth Century*. London: Longman.

Bauer, L. 2003 (1993). *Introducing Linguistic Morphology*. Edinburgh: Edinburgh University Press.

Bauer, L. 2004. *A Glossary of Morphology*. Edinburgh: Edinburgh University Press.

Bauer, L. 2007. *The Linguistics Student’s Handbook*. Edinburgh: Edinburgh University Press.

Bauer, L. 2015. ‘Australian and New Zealand Englishes’. In M. Filppula, J. Klemola & D. Sharma (eds), *The Oxford Handbook of World Englishes* (pp. 409–24).Oxford: Oxford University Press.

Bauer, L. & Bauer, W. 2002. ‘Can we watch regional dialects developing in colonial English? The case of New Zealand’. *English World Wide*, **23**(2), 169–93.

Bauer, L. & Bauer, W. 2012. ‘The inflection-derivation divide in Māori and its implications’. *Te Reo*, **55**, 3–24.

Bauer, L. & Warren, P. 2004. ‘New Zealand English: Phonology’. In B. Kortmann, E. Schneider, K. Burridge, R. Mesthrie & C. Upton (eds), *A Handbook of Varieties of English* (pp. 580–602). Berlin: Mouton de Gruyter.

Bauer, L., Warren, P., Bardsley, D., Kennedy, M. & Major, G. 2007. ‘Illustrations of the IPA: New Zealand English’. *Journal of the International Phonetic Association*, **37**(1), 97–102.

Bauer, W.2001. *Māori: Descriptive Grammars*. Abingdon: Routledge.

Bauer, W. 2003 [1993]. *Māori*. Abingdon: Routledge.

Bell, A. 1984. ‘Language style as audience design’. *Language in Society*, **13**, 145–204.

Bell, A. 2014. *The Guidebook to Sociolinguistics*. Oxford: Wiley Blackwell.

Bell, A. & Gibson, A. 2008. Stopping and fronting in New Zealand Pasifika English. *University of Pennsylvania Working Papers in Linguistics: A Selection of Papers from NWAV*, **36**, 14(2), 42–53.

Bell, A. & Kuiper, K. (eds) 2001. *New Zealand English*. Wellington, NZ: Victoria

University Press.

Benczes, R. 2019. *Rhyme over Reason: Phonological Motivation in English*. Cambridge: Cambridge University Press.

Berko, J. 1958. 'The child's learning of English morphology'. *Word*, **14**, 150–77.

Berko, J. & Brown, R. 1960. 'Psycholinguistic research methods'. In P. Mussen (ed.), *Handbook of Research Methods in Child Development*. New York: John Wiley.

Berko Gleason, J. 2005. *The Development of Language*. Boston: Pearson.

Biber, D. 1988. *Variation Across Speech and Writing*. New York: Cambridge University Press.

Blake, B. 2007. *Playing with Words: Humour in the English Language*. London: Equinox.

Blake, B. 2008. *All About Language*. Oxford: Oxford University Press.

Bloomfield, L. 1933. *Language*. Chicago: University of Chicago Press.

Bolinger, D.W. 1949. 'The sign is not arbitrary'. *Boletín del Instituto Caro y Cuervo* (thesaurus), **5**, 52–62.

Bolinger, D.W. 1975. *Aspects of Language*. New York: Harcourt Brace Jovanovich.

Bolinger, D.W. 1980. *Language: The Loaded Weapon*. London: Longman.

Bonnet, A & Siemund, P. (eds) 2018. *Foreign Language Education in Multilingual Classrooms*. Amsterdam: Walter Benjamins.

Börjars, K. & Burridge, K. 2019. *Introducing English Grammar*. 3rd edn. London: Routledge.

Bourdieu, P. 1991. *Language and Symbolic Power*. Ed. by J.B. Thompson; trans. by G. Raymond & M. Adamson. Cambridge: Polity Press.

Bowern, C. 2012. 'The riddle of Tasmanian languages'. *Proceedings of the Royal Society of London B: Biological Sciences*, **279**(1747), 4590–5.

Bradley, D. 2004. 'Regional characteristics of Australian English phonology'. In K. Burridge & B. Kortmann (eds), *Varieties of English: The Pacific and Australasia*, Vol. **3** (pp. 111–23). Berlin: Mouton de Gruyter.

Britain, D. 1992. 'Linguistic change in intonation: The use of high rising terminals in New Zealand English'. *Language Variation and Change*, **4**, 77–104.

Brown, E.K. & Miller, J. 2002. *Syntax: A Linguistic Introduction to Sentence Structure.* London: Routledge.

Brown, K. (ed.) 2006. *Encyclopedia of Language and Linguistics* [electronic resource]. 2nd edn. Boston: Elsevier.

Brown, P. & Levinson, S. 1987. *Politeness: Some Universals in Language*. Cambridge: Cambridge University Press.

Brown, R. 1973. *A First Language*. Sydney: Allen & Unwin.

Bryant, P. 1997. 'A dialect survey of the lexicon of Australian English'. *English World-Wide*, **18**, 211–41.

Burdelski, M. 2013. 'Socializing children to honorifics in Japanese: Identity and stance in interaction'. *Multilingua*, **32**(2), 247–73.

Burnham, D. & Kruger, P. 2011. 'Capturing the Aussie accent'. *AM*, ABC Radio, 26 January, http://www.abc.net.au/am/content/2011/s3121938.htm.

Burridge, K. 2002. 'Changes within Pennsylvania German grammar as enactments of anabaptist world-view'. In N. Enfield (ed), *Ethnosyntax: Explorations in Grammar and Culture* (pp. 207–30). Oxford: Oxford University Press.

Burridge, K. 2007. 'Language contact and convergence in Pennsylvania German'. In A. Aikhenvald & R. Dixon (eds), *Grammars in Contact: A Cross-linguistic Typology: Explorations in Linguistic Typology*, Vol. **4** (pp. 179–200). Oxford: Oxford University Press.

Burridge, K. 2010. 'Linguistic cleanliness is next to godliness: Taboo and purism'. *English Today*, **26**(2): 3–13.

Burridge, K. 2012. 'Euphemism and language change', *Lexis, E-Journal in English Lexicology*, https://journals.openedition.org/lexis/355.

Burridge, K. & Florey, M. 2002. '"Yeah-no he's a good kid": A discourse analysis of *yeah-no* in Australian English', *The Australian Journal of Linguistics*, **22**(2),149–72.

Butcher, A. 2008. 'Linguistic aspects of Australian Aboriginal English'. *Clinical Linguistics and Phonetics*, **22**(8), 625–42.

Butler, J. 1999. *Gender Trouble: Feminism and the Subversion of Identity*. New York: Routledge.

Butters, R.R. 2007. 'A linguistic look at trademark dilution'. *Santa Clara High Technology Law Journal*, **24**(3), 507.

Butters, R.R. 2010. 'Trademark linguistics – trademarks: Language that one owns'. In M. Coulthard & A. Johnson (eds), *The Routledge Handbook of Forensic Linguistics* (pp. 351–64). Abingdon: Routledge.

Bybee, J. 2003a. 'Mechanisms of change in grammaticization: The role of frequency'. In B. Joseph & R. Janda (eds), *The Handbook of Historical Linguistics* (pp. 602–23). Oxford: Blackwell.

Bybee, J. 2003b. *Phonology and Language Use*. Cambridge: Cambridge University Press.

Cai, P. 2013. ‘Business can benefit from students of humanities’. *The Age*, 28 January.

Calude, A. 2017. ‘The use of heaps as quantifier and intensifier in New Zealand English’ *English Language and Linguistics*, December, doi: 10.1017/S1360674317000521.

Cameron, D. 1997. ‘Reply to James Milroy’. *Journal of Sociolinguistics*. **1**(1), 163–5.

Cameron, D. 1999. ‘Performing gender identity: Young men’s talk and the construction of heterosexual masculinity’. In A. Jaworski & N. Coupland (eds), *The Discourse Reader*, 2nd edn (pp. 419–32). London: Routledge.

Cameron, D. 2012. *Verbal Hygiene*. 2nd edn. London: Routledge.

Cameron, D., Frazer, E. Harvey, P., Rampton, M.B.H. & Richardson, K. 1992. *Researching Language: Issues of Power and Method*. London: Routledge.

Campbell, L. 2004. *Historical Linguistics: An introduction*. 2nd edn. Edinburgh: Edinburgh University Press.

Campbell, L. & Janda, R. (eds) 2001. *Language Sciences*, **23** (special edition on grammaticalization theory).

Carpenter, H. (ed.) 1981. *The Letters of J.R.R. Tolkien* (with the assistance of Christopher Tolkien). London: Allen & Unwin.

Carr, P. 1993. *Phonology*. London: Macmillan.

Carstairs-McCarthy, A. 2002. *An Introduction to English Morphology*. Edinburgh: Edinburgh University Press.

Catford, J. 2001. *A Practical Introduction to Phonetics.* 2nd edn. New York: Oxford University Press.

Cazden, C. 1972. *Child Language and Education*. New York: Holt, Rhinehart and Winston.

Chafe, W. 1994. *Discourse, Consciousness, and Time: The Flow and Displacement of Conscious Experience in Speaking and Writing*. Chicago: University of Chicago Press.

Chomsky, N. 1965. *Aspects of the Theory of Syntax.* Cambridge, MA: MIT Press.

Clark, J., Yallop, C. & Fletcher, J. 2007. *An Introduction to Phonetics and Phonology*. 3rd edn. Oxford: Wiley-Blackwell.

Clarke, A.C. 1972. *The Lost Worlds of 2001: Science Fiction*. New York: Signet.

Cleland Silva, T. 2018. *Transnational Management and Globalised Workers: Nurses Beyond Human Resources*. London: Routledge.

Clyne, M. 2003. *Dynamics of Language Contact*. Cambridge: Cambridge University Press.

Clyne, M. 2008. 'The monolingual mindset as an impediment to the development of plurilingual potential in Australia'. *Sociolinguistic Studies*, **2**(3), 347–66.

Coates, J. 2014. *Women, Men and Language*. 3rd edn. London: Longman.

Cohen, R. 1995. 'There's more to a name'. *Stanford Business School Magazine,* March, 8–9.

Comrie, B. 1979. 'Russian'. In T. Shopen (ed.), *Languages and Their Status* (pp. 91–150). Cambridge, MA: Winthrop.

Comrie, B. 1987. *The World's Major Languages*. London: Croom Helm.

Cooke, M. 1995. 'Aboriginal evidence in the cross-cultural courtroom'. In D. Eades (ed.), *Language in Evidence: Issues Confronting Aboriginal and Multicultural Australia* (pp. 55–96). Sydney: UNSW Press.

Cordella, M. 2004. *The Dynamic Consultation: A Discourse Analytical Study of Doctor–Patient Communication*. Amsterdam: John Benjamins.

Cordella, M. & Musgrave, S. 2009. 'Oral communication skills of international medical graduates: Assessing empathy in discourse'. *Communication and Medicine*, **6**(2), 129–42.

Coulmas, F. 2002. *Writing Systems: An Introduction to Their Linguistic Analysis*. Cambridge: Cambridge University Press.

Coulthard, M. & Johnson, A. 2007. *An Introduction to Forensic Linguistics: Language in Evidence*. London: Routledge.

Coupland, N., Coupland, J., Giles, H. & Henwood, K. 1988. 'Accommodating the elderly: Invoking and extending a theory'. *Language in Society*, **17**(1): 1–41.

Cox, F. 2006. 'Australian English pronunciation into the 21st century'. *Prospect: Australian Journal of TESOL*, **21**, 3–21.

Cox, F. 2012. *Australian English Pronunciation and Transcription*. Melbourne: Cambridge University Press.

Cox, F. & Palethorphe, S. 2007. 'Illustrations of the IPA: Australian English'. *Journal of the International Phonetics Association*, **37**(3), 341–50.

Crowley, T. & Bowern, C. 2010. *An Introduction to Historical Linguistics*. Oxford: Oxford University Press.

Crowley, T., Lynch, J., Siegel, J. & Piau, P. 1995. *The Design of Language: An Introduction to Descriptive Linguistics*. Auckland: Longman Paul.

Cruse, A. 2004. *Meaning in Language*. Oxford: Oxford University Press.

Crystal, D. 1994. 'Documenting rhythmical change'. In J. Windsor Lewis (ed.) *Studies in general and English phonetics* (pp. 174–9). London: Routledge.

Crystal, D. 1997. *The Cambridge Encyclopedia of Language*. 2nd edn. Cambridge: Cambridge University Press.

Crystal, D. 2004. *The Stories of English*. Harmondsworth: Penguin.

Culpepper, J. 2011. *Impoliteness: Using Language to Cause Offence*. Cambridge: Cambridge University Press.

Cushing, S. 1994. *Fatal Words: Communication Clashes and Airport Crashes*. Chicago: University of Chicago Press.

Czerniewska, P. 1985. 'The experience of writing'. In *Open University, Every Child's Language: An In-service Pack for Primary Teachers*. Clevedon: Multilingual Matters.

Damico, J., Müller, N. & Ball, M. 2010. *Handbook of Language and Speech Disorders*. Hoboken, NJ: John Wiley and Sons.

Damousi, J. 2010. *Colonial Voices*. Cambridge: Cambridge University Press.

De Beaugrande, R. & Dressler, W. 1996. *Introduction to Text Linguistics*. New York: Routledge.

de Saussure, F. 1966. *Course in General Linguistics*. Eds C. Bally & A. Sechehaye in collaboration with A. Riedlinger. Trans. W. Baskin. New York: McGraw-Hill.

Degani, M. & Onysko, A. 2010. 'Hybrid compounding in New Zealand English: Report'. *World Englishes*, **29**(2), 209–33.

Dennett, D.C. 2013. *Intuition Pumps and Other Tools for Thinking*. New York: W.W. Norton.

Deutscher, G. 2010. *Through the Looking Glass: Why the World Looks Different in Other Languages*. New York: Metropolitan Books.

Dewaele, J.M. 2010. *Emotions in Multiple Languages*. Basingstoke: Palgrave Macmillan.

Dickson, G. 2016. 'Rethinking the substrate languages of Roper Kriol: The case of Marra'. In F. Meakins & C. O'Shannessy (eds), *Loss and Renewal: Australian Languages Since Colonisation* (pp. 145–73). Berlin: De Gruyter Mouton.

Dixon, R.M.W. & Aikhenvald, A.Y. (eds) 2007. *Word: A Cross-linguistic Typology*. Cambridge: Cambridge University Press.

Domingue, N.Z. 1977. ‘Middle English: Another creole?’ *Journal of Creole Studies*, **1**, 89–108.

Dorian, N.C. 1994. ‘Purism vs compromise in language revitalization and language revival’. *Language in Society*, **23**(4), 479–94.

Douglas, M. 1966. *Purity and Danger: An Analysis of Concepts of Pollution and Taboo*. London: Routledge & Kegan Paul.

Douglas, S. & Stirling, L. 2012. ‘Metacommunication, social pretend play and children with autism’. *Australasian Journal of Early Childhood*, **37**(4), 34–43.

Douglas, S. & Stirling, L. (eds) 2016. *Children’s Play, Pretence, and Story: Studies in Culture, Context and ASD.* London: Routledge.

Dronkers, N.F., Plaisant, O., Iba-Zizen, M.T. & Cabanis, E.A. 2007. ‘Paul Broca’s historic cases: High resolution MR imaging of the brains of Leborgne and Lelong’. *Brain*, **130**(5), 1432–41.

Du Ponceau, P.S. 1819. *Report of the Corresponding Secretary to the Committee, of His Progress in the Investigation Committed to Him of the General Character and Forms of the Languages of the American Indians*, 12 January, **1**, xvii–xlvi.

Dubinsky, S. & Holcomb, C. 2011. *Understanding Language Through Humor*. Cambridge: Cambridge University Press.

Eades, D. 2010. *Sociolinguistics and the Legal Process*. Bristol: Multilingual Matters.

Eades, D. 2012. ‘Communication with Aboriginal speakers of English in the legal process’. *Australian Journal of Linguistics*, **32**(4),473–89.

Eades, D. 2013. *Aboriginal Ways of Using English*. Canberra: Aboriginal Studies Press.

Eagleson, R. 1994. ‘Forensic analysis of personal written texts: A case study’. In J. Gibbons (ed.), *Language and the Law* (pp. 362–73). London: Longman.

Easton, C. 2008. ‘Discourses of orthography development: Community-based practice in Milne Bay PNG’. Unpublished PhD thesis, Charles Sturt University, Bathurst.

Evans, C. 2013. ‘Forensic linguistics: Top ten trademark cases’. Wordpress, http://uwelingo.wordpress.com/2013/03/03/forensic-linguistics-top-ten-trademark-cases.

Evans, N. 2010. *Dying words: Endangered Languages and What They Have to Tell Us*. Malden, MA: Wiley-Blackwell.

Feagin, C. 2004. ‘Entering the community: Fieldwork’. In J. Chambers, P. Trudgill & N. Schilling-Estes (eds), *Handbook of Language Variation and Change*. London: Blackwell.

Ferguson, C. 1959. 'Diglossia'. *Word*, **15**, 325–40.

Ferguson, N. 2008. 'The Americanisation of Australian English: Attitudes, perception and usage', BA Hons thesis, Monash University, Melbourne.

Field, J. 2005. *Language and the Mind*. London: Routledge.

Filppula, M. 2008. 'Irish English: Morphology and syntax'. In K. Berndt & C. Upton (eds), *The British Islands*, Vol. **1** (pp. 328–59). Berlin: Walter de Gruyter.

Finch, G. 2003. *How To Study Linguistics: Guide to Understanding Language*. Basingstoke: Palgrave Macmillan.

Finkelstein, S.R. 2019. 'Swearing and the brain'. In K. Allan (ed.), *The Oxford Handbook of Taboo Words and Language* (pp. 108–39). Oxford: Oxford University Press.

Firth, J. R. 1957. 'A synopsis of linguistic theory, 1930–1955'. In J. Firth et al. *Studies in Linguistic Analysis: Special Volume of the Philological Society*. Oxford: Blackwell.

Fletcher, J. & Loakes, D. 2006. 'Patterns of rising and falling in Australian English'. In P. Warren & C.I. Watson (eds), *Proceedings of the 11th Australian International Conference on Speech Science and Technology 2006* (pp. 42–47).

Foster-Cohen, S. 1999. *An Introduction to Child Language Development*. London: Addison Wesley

Fraser, H. 2018. 'Assisting hearers to hear words that aren't there: Dangers in using police transcripts of indistinct covert recordings'. *Australian Journal of Forensic Sciences* **50**(2), 129–39.

Frellesvig, B. 2010. *A History of the Japanese Language*. Cambridge: Cambridge University Press.

Fujinami, Y. 1996. 'An implementation of Japanese grammar based on HPSG'. Unpublished MSc thesis, University of Edinburgh.

Gardner, H. 1974. *The Shattered Mind: The Person after Brain Damage*. New York: Vintage Books.

Garner, J.F. 1994. *Politically Correct Bedtime Stories: Modern Tales for Our Life and Times.* London: Macmillan.

Garrod, S. 2006. 'Psycholinguistic research methods'. In R. Asher & J. Simpson (eds), *Encyclopedia of Language and Linguistics*. 2nd edn. (pp. 251–7). London: Pergamon Press.

Gawne, L. 2011. 'Gillard, grammar and the language of politics'. *Crikey*, 15 September, http://blogs.crikey.com.au/fullysic/2011/09/15/gillard-grammar-and-the-language-of-politics.

Gee, J. 1990. *An Introduction to Discourse Analysis: Theory and Method.* London: Routledge.

Ghomeshi, J., Jackendoff, R., Rosen, N. & Russell, K. 2004. 'Contrastive focus reduplication in English (the SALAD-salad paper)'. *Natural Language and Linguistic Theory*, **22**(2), 307–57.

Giles, H. & Powesland, P. 1975. *Speech Style and Social Evaluation*. London: Academic Press.

Giles, H. & Smith, P. 1979. 'Accommodation theory: Optimal levels of convergence'. In H. Giles & R.N. St Clair (eds), *Language and Social Psychology* (pp. 45–65). Baltimore, MD: Basil Blackwell.

Goddard, C. (ed.) 2008. *Cross-linguistic Semantics*. Amsterdam: John Benjamins.

Goddard, C. (ed.) 2011. *Semantic Analysis: A Practical Introduction*. 2nd edn. Oxford: Oxford University Press.

Goddard, C. & Wierzbicka, A. (eds) 2002. *Meaning and Universal Grammar: Theory and Empirical Findings*. Amsterdam: John Benjamins.

Goffman, E. 1967. *Interaction Ritual*. Chicago: Aldine.

Gottlieb, N. 2005. *Language and Society in Japan*. Cambridge: Cambridge University Press.

Graddol, D., Cheshire, J. & Swann, J. 1987. *Describing Language*. Milton Keynes: Open University Press.

Grant, T. 2013. 'The linguistic clues that reveal your true Twitter identity'. *The Conversation*, 20 November, https://theconversation.com/the-linguistic-clues-that-reveal-your-true-twitter-identity-2045.

Greenberg, J. 1960. 'A quantitative approach to the morphological typology of language'. *International Journal of American Linguistics*, **26**(3), 178–84.

Grice, P.H. 1975. 'Logic and conversation'. In P. Cole & J. Morgan (eds), *Syntax and Semantics*. Vol. **3** (pp. 41–58). New York: Academic Press.

Halliday, M.A.K. & Hasan, R. 1976. *Cohesion in English*. London: Longman.

Hamers, J.F. & Blanc, M.H.A. 2000. *Bilinguality and Bilingualism*. 2nd edn. Cambridge: Cambridge University Press.

Harlow, R. 2007. *Maori: A Linguistic Introduction*. Cambridge: Cambridge University Press.

Harrington, J., Cox, F. & Evans, Z. 1997. 'An acoustic study of broad, general and cultivated Australian English vowels'. *Australian Journal of Linguistics*, **17**: 155–84.

Harrington, J., Palethorpe, S. & Watson, C.I. 2000. ‘Monophthongal vowel changes in Received Pronunciation: An acoustic analysis of the Queen’s Christmas broadcasts’. *Journal of the International Phonetic Association*, **30**(1/2), 163–78.

Harrington, J., Palethorpe, S. & Watson, C.I. 2005. ‘Deepening or lessening the divide between diphthongs: An analysis of the Queen’s annual Christmas broadcasts’. In W.J. Hardcastle & J.M. Beck (eds), *A Figure of Speech: A Festschrift for John Laver* (pp. 227–62). Mahwah, NJ: Lawrence Erlbaum.

Hartnett, M. 1975. *A Farewell to English, and Other Poems*. Dublin: Gallery Books.

Haspelmath, M. 2002. *Understanding Morphology*. Oxford: Oxford University Press.

Haspelmath, M. & Sims, A.D. 2010. *Understanding Morphology*. London: Hodder.

Haugen, E. 1966. ‘Dialect, language, nation’. *American Anthropologist*, **68**, 922–35.

Hay, J. & Drager, K. 2010. ‘Stuffed toys and speech perception’. *Linguistics*, **48**(4), 865–92.

Hay, J., Maclagan, M. & Gordon, E. 2008. *New Zealand English*. Edinburgh: Edinburgh University Press.

Heath, D. & Heath, C. 2011. ‘How to pick the perfect brand name’. *Fast Company*, http://www.fastcompany.com/1702256/how-pick-perfect-brand-name.

Heigham, J. & Croker, R. 2009. *Qualitative Research in Applied Linguistics: A Practical Introduction*. Basingstoke: Palgrave Macmillan.

Hendery, R. 2012. ‘The relationship between language ideology and language change in a small, isolated community: The case of Palmerston Island’. *Te Reo: Journal of the Linguistics Society of New Zealand*, **55**, 25–49.

Hinton, L., Nichols, J. & Ohala, J.J. 1994. *Sound Symbolism*. Cambridge: Cambridge University Press.

Hochfeder, D. 2012. *The Telegraph in America, 1832–1920*. Baltimore, MD: Johns Hopkins University Press.

Hock, H.H. & Joseph, B.D. 1996. *Language History, Language Change, and Language Relationship: An Introduction to Historical and Comparative Linguistics*. Berlin: Mouton de Gruyter.

Hockett, C.F. 1960. ‘The origin of speech’. *Scientific American*, **203**, 88–96.

Hockett, C.F. & Altmann, S. 1968. ‘A note on design features’. In T.A. Sebeok (ed.), *Animal Communication: Techniques of Study and Results of Research* (pp. 61–72). Bloomington, IN: Indiana University Press.

Hodge, G. & Johnston, T. 2014. ‘Points, depictions, gestures and enactment: Partly lexical and non-lexical signs as core elements of single clause-like units in Auslan (Australian Sign Language)’. *Australian Journal of Linguistics*, **34**(2), 262–91.

Hofstadter, D.R. 1997. *Le Ton Beau de Marot: In Praise of the Music of Language*. New York: Basic Books.

Hofstadter, D.R. & Dennett, D.C. 1982. *The Mind’s I: Fantasies and Reflections on Self and Soul*. New York: Bantam Books.

Hofstede, G. 1980. *Culture’s Consequences: International Differences in Work-related Values*. Beverly Hills, CA: Sage.

Holldobler, B. & Wilson, E.O. 1977. ‘The multiple recruitment systems of the African weaver ant *Oecophylla longinoda* (Latreille) (Hymenopetera: Formicidae)’. *Behavioral Ecology and Sociobiology*, **3**, 19–60.

Holmes, J. 1988. ‘Of course: A pragmatic particle in New Zealand women’s and men’s speech’. *Australian Journal of Linguistics*, **8**(1), 49–74.

Holmes, J. 1994. ‘New Zealand flappers: An analysis of T voicing in New Zealand English’. *English World-wide*, **15**, 195–224.

Holmes, J. 1995. *Women, Men and Politeness*. London: Longman.

Hopper, P.J. & Traugott, E C. 2003. *Grammaticalization*. 2nd edn. Cambridge: Cambridge University Press.

Horvath, B. 1985. *Variation in Australian English: The Sociolects of Sydney*. Cambridge: Cambridge University Press.

Householder, F.W. 1971. *Linguistic Speculations*. Cambridge: Cambridge University Press.

Huddleston, R. & Pullum, G.K. 2005. *A Student’s Introduction to English Grammar*. Cambridge: Cambridge University Press.

Hudson, R. 1984. *Higher-level Differences between Speech and Writing*. London: Committee for Linguistics in Education.

Hudson, R. 2001. ‘Grammar teaching and writing skills: The research evidence’. *Syntax in the Schools*, **17**, 1–6.

Hudson, R. 2004. ‘Why education needs linguistics (and vice versa)’. *Journal of Linguistics*, **40**, 105–30.

Hudson, R. 2007. ‘How can key skills “sell” linguistics to students and employers’, http://dickhudson.com/papers.

Hughes G. 1991. *Swearing: A Social History of Foul Language, Oaths and Profanity in English*. Oxford: Blackwell.

Hurst, D. 2014. ‘Mathias Cormann denies his “economic girlie-man” insult is sexist’. *The Guardian* (Australia), 19 October, http://www.theguardian.com/australia-news/2014/oct/19/mathias-cormann-denies-his-economic-girlie-man-insult-is-sexist.

Huth, A.G., de Heer, W.A., Griffiths, T.L. Theunissen, F.E. & Gallant, J.L. 2016. ‘Natural speech reveals the semantic maps that tile human cerebral cortex’. *Nature*, **532**, 453–8.

Hymes, D. 1972. Models of the interaction of language and social life. In J. Gumperz & D. Hymes (eds), *Directions in Sociolinguistics: The Ethnography of Communication* (pp. 35–71). New York: Holt, Rhinehart & Winston.

Indefrey, P. & Levelt, W.J.M. 2004. ‘The spatial and temporal signatures of word production components’. *Cognition*, **92**, 101–44.

International Phonetic Association 1999. *Handbook of the International Phonetic Association.* Cambridge: Cambridge University Press.

Ishida, K. 2006. ‘How can you be so certain? The use of hearsay evidentials by English-speaking learners of Japanese’. *Journal of Pragmatics,* **38**, 1281–1304.

Iwasaki, S., Bartlett, M., Manns, H. & Willoughby, L. 2018. The challenges of multimodality and multi-sensoriality: Methodological issues in analyzing tactile signed interaction. *Journal of Pragmatics,* doi:10.1016/j.pragma.2018.05.003.

Jay, T. 2000. *Why We Curse: A Neuro-Psycho-Social Theory of Speech.* Philadelphia, PA: John Benjamins.

Jernudd, B. 1989. ‘The texture of language purism: an introduction’. In B.H. Jernudd & M.J. Shapiro (eds), *The Politics of Language Purism* (pp. 1–20). Berlin: Mouton de Gruyter.

Johnson, S. 1755. *Dictionary of the English Language*. London: W. Strahan.

Johnston, T. & Schembri, A. 2007. *Australian Sign Language (Auslan): An Introduction to Sign Language Linguistics*. Cambridge: Cambridge University Press.

Jones, C. & Meakins, F. 2013. ‘The phonological forms and perceived functions of janyarrp, the Gurindji “baby talk” register’. *Lingua*, **134**, 170–93.

Joos, M. 1961. *The Five Clocks*. New York: Harcourt, Brace and World.

Jurafsky, D. & Martin, J.H. 2008. *Speech and Language Processing.* 2nd edn. Upper Saddle River, NJ: Prentice Hall.

Kanno, Y. 2004. ‘Sending mixed messages: Language minority education at a Japanese public elementary school’. In A. Pavlenko & A. Blackledge (eds), *Negotiation of Identities in Multilingual Contexts* (pp. 316–38). Clevedon: Multilingual Matters.

Katamba, F. 1994. *English Words: Structure, History and Usage*. London: Routledge.

Kelly, B.F., Leben, W.R. & Cohen, R.H. 2003. ‘The meaning of consonants’. In *Proceedings of the Twenty-Ninth Annual Meeting of the Berkeley Linguistics Society: General Session and Parasession on Phonetic Sources of Phonological Patterns: Synchronic and Diachronic Explanations* (pp. 245–53).

Kendon, A. 2004. *Gesture: Visible Action as Utterance*. Cambridge: Cambridge University Press.

Kenneally, C. 2007. *The First Word: The Search for the Origins of Language*. New York: Viking.

Kennedy, G. (ed.) 1999. *New Zealand Sign Language: Distribution, Origins, Reference*. Wellington: Deaf Studies Research Unit, Victoria University.

Kennedy, G., Arnold, R., Dugdale, P., Fahey, S. & Moskovitz, D. (eds) 1997. *A Dictionary of New Zealand Sign Language*. Auckland: Auckland University Press with Bridget Williams Books.

Klink, R. 2000. 'Creating brand names with meaning: The use of sound symbolism'. *Marketing Letters*, **11**(1), 5–20.

Knecht, S., Dräger, B., Flöel, A., Lohmann, H., Breitenstein, C., Deppe, M., Henningsen, H. & Ringelstein, E.-B. 2001. 'Behavioural relevance of atypical language lateralization in healthy subjects'. *Brain: A Journal of Neurology*, **124**(8), 1657–65.

Koch, H. 1985. 'Non-standard English in an Aboriginal land claim'. In. J.B. Pride (ed.), *Cross-cultural Encounters: Communication and Miscommunication* (pp. 176–95). Melbourne: River Seine.

Konishi, M. 2010. 'From central pattern generator to sensory template in the evolution of birdsong'. *Brain and Language*, **15**, 18–20.

Koteyko, N. 2014. *Language and Politics in Post-Soviet Russia: A Corpus Assisted Approach*. London: Palgrave Macmillan.

Kruger, H. & Smith, A. 2018. 'Colloquialization and densification in Australian English: A multidimensional analysis of the Australian Diachronic Hansard Corpus'. *Australian Journal of Linguistics*, **38**(3): 293–328.

Kubota, R. & McKay, S. 2009. 'Globalization and language learning in rural Japan: The role of English in the local linguistic ecology'. *TESOL Quarterly*, **43**, 593–619.

Labov, W. 1963. 'Social motivations of a sound change'. *Word*, **19**, 273–309.

Labov, W. 1966. *The Social Stratification of English in New York City*. Washington, DC: Centre for Applied Linguistics.

Labov, W. 1969. 'The logic of non-standard English'. *Georgetown Monograph on Language and Linguistics*, **22**, 1–44.

Labov, W. 1972a. *Language in the Inner City: Studies in Black English Vernacular*. Philadelphia, PA: University of Pennsylvania Press.

Labov, W. 1972b. *Sociolinguistic Patterns*. Oxford: Basil Blackwell.

Labov, W. 1972c. 'Some principles of linguistic methodology'. *Language in Society*, **1**, 97–120.

Labov, W. 1997. 'How I got into linguistics, and what I got out of it', http://www.ling.upenn.edu/~wlabov/HowIgot.html.

Labov, W. 2007. 'Transmission and diffusion'. *Language: Journal of the Linguistic Society of America*, **83**(2), 344–87.

Labov, W. & Waletzky, J. 1967. 'Narrative analysis'. In J. Helm (ed.), *Essays on the Verbal and Visual Arts* (pp. 12–44). Seattle: University of Washington Press.

Labrune, L. 2012. *The Phonology of Japanese*. New York: Oxford University Press.

Labrune, L. 2014. 'The phonology of Japanese /r/: A panchronic account'. *Journal of East Asian Linguistics*, **23**(1), 1–25.

Lacey, R. 1987. *Ford: The Men and the Machine*. London: William Heinemann.

Ladefoged, P. & Disner, S.F. 2012. *Vowels and Consonants*. 3rd edn. Oxford: Wiley-Blackwell.

Ladefoged, P. & Johnson, K. 2011. *A Course in Phonetics.* 6th edn. Boston: Cengage.

Ladefoged, P. & Maddieson, I. 1996. *The Sounds of the World's Languages*. Oxford: Blackwell.

Lakoff, G. & Johnson, M. 1981. *Metaphors We Live By*. Chicago: University of Chicago Press.

Langacker, R.W. 1977. 'Syntactic reanalysis'. In C.N. Li (ed.), *Mechanisms of Syntactic Change* (pp. 59–139). Austin, TX: University of Texas Press.

Langford, I. 2000. 'Forensic semantics: The meaning of murder, manslaughter and homicide'. *Forensic Linguistics*, **7**(1), 72–94.

LaPolla, R.J. 1994. 'An experimental investigation into phonetic symbolism as it relates to Mandarin Chinese'. In L. Hinton, J. Nichols & J.J. Ohala (eds), *Sound Symbolism* (pp. 130–47). Cambridge: Cambridge University Press.

Larson, M.L. 1998. *Meaning-based Translation: A Guide to Cross Language Equivalence*. 2nd edn. Lanham, MD: University Press of America.

Lass, R. 2000. *Phonology: An Introduction to Basic Concepts*. Cambridge: Cambridge University Press.

Leech, G. 1983. *Principles of Pragmatics*. London: Longman.

Lentine, G. & Shuy, R.W. 1989. '*Mc*-meaning in the market place'. *American Speech*, **65**(4), 349–66.

Liddicoat, A.C. 2000–01. 'Learning a language, learning about language, learning to be literate'. *Babel*, **35**(3), 12–16.

Litosseliti, L. (ed.) 2010. *Research Methods in Linguistics*. London: Continuum.

Liu, W. & Kroodsma, D.E. 1999. 'Song development by chipping sparrows and field sparrows'. *Animal Behaviour*, **57**, 1275–86.

Loakes, D., Clothier, J.J., Hajek, J. & Fletcher, J. 2014. ‘An investigation of the /el/-/ael/ merger in Australian English: A pilot study on production and perception in south-west Victoria’. *Australian Journal of Linguistics*, **34**(4), 436–52.

Loftus, E.F. 1975. ‘Leading questions and the eyewitness report’. *Cognitive Psychology*, **7**, 560–72.

Loftus, E.F. 1979. *Eyewitness Testimony*. Cambridge, MA: Harvard University Press.

Lowrey, T.M. & Shrum, L.J. 2007. *Journal of Consumer Research*, **34**(3), 406–14.

Lyons, J. 1968. *Introduction to Theoretical Linguistics*. Cambridge: Cambridge University Press.

Lyons, J.1977. *Semantics: Volume 1*. Cambridge: Cambridge University Press.

Lyons, J. 1995 *Linguistic Semantics*. Cambridge: Cambridge University Press.

MacKay, D.G., Shafto, M., Taylor, J.K., Marian, D., Abrams, L. & Dyer, J. 2004. ‘Relations between emotion, memory and attention: Evidence from taboo Stroop, lexical decision, and immediate memory tasks’. *Memory and Cognition*, **32**, 474–88.

Mair, V. 2013. ‘Too many English loanwords in Japanese’. *Language Log*, 12 July, http://languagelog.ldc.upenn.edu/nll/?p=5176.

Maitz, P. & Volker, C.A. 2017. Documenting Unserdeutsch: Reversing colonial amnesia. *Journal of Pidgin and Creole Languages* **32**(2), 365–97.

Makino, S. & Tsutsui, M. 1986. *A Dictionary of Basic Japanese Grammar*. Tokyo: Japan Times.

Malcolm, I.G. 2004a. ‘Australian creoles and Aboriginal English: Morphology and syntax’. In B. Kortmann, K. Burridge, R. Mesthrie, EW. Schneider & C. Upton (eds), *A Handbook of Varieties of English*, Vol. **2** (pp. 657–81). Berlin: Mouton de Gruyter.

Malcolm, I.G. 2004b. ‘Australian creoles and Aboriginal English: Phonology’. In E.W. Schneider, K. Burridge, B. Kortmann, R. Mesthrie & C. Upton (eds), *A Handbook of Varieties of English*, Vol. **1** (pp. 656–70). Berlin: Mouton de Gruyter.

Mannell, R., Cox, F. & Harrington, J. 2014. *An Introduction to Phonetics and Phonology*. Sydney: Macquarie University.

Manns, H. & Burridge, K. 2016. ‘Hustings and human speech (failings) in a 24/7 campaign’. *The Conversation*, 15 June, https://theconversation.com/hustings-and-human-speech-failings-in-a-24-7-campaign-61088

Marmion, D., Obata, K. & Troy, J. 2014. *Community, Identity, Wellbeing: The Report of the Second National Indigenous Languages Survey*. Canberra: Australian

Institute of Aboriginal and Torres Strait Islander Studies.

Marsden, S. 2007. 'On the emergence of regional varieties of New Zealand English'. *New Zealand English Journal*, **21**, 64–72.

Matsumoto, Y. 1988. 'Reexamination of the universality of face: Politeness phenomena in Japanese'. *Journal of Pragmatics*, **12**(4), 403–26.

Mauthner, F. 1982 [1901]. *Beiträge ze einer Kritik der Sprache*. [3 vols]. Frankfurt: Ullstein.

Mayer, M. 1969. *Frog, Where are You?* New York: The Dial Press.

McBride C. 2015. *Children's Literacy Development: A Cross-Cultural Perspective on Learning to Read and Write*. London: Routledge.

McBride C. 2016. Is Chinese special? Four aspects of Chinese literacy acquisition that might distinguish learning Chinese from learning alphabetic orthographies. *Educational Psychology Review* **28**, 523–49.

McColl, R.M. 2007. *Trask's Historical Linguistics*. 2nd edn. London: Hodder Arnold.

McCulloch, G. 2019. *Because Internet: Understanding the New Rules of Language Online*. New York: Riverhead (Penguin).

McCulloch, G. & Gawne, L. 2018. 'Emoji grammar as beat gestures'. *CEUR Workshop Proceedings*, http://ceur-ws.org/Vol-2130/short1.pdf.

McEnery, T. & Hardie, A. 2011. *Corpus Linguistics: Method, Theory and Practice*. Cambridge: Cambridge University Press.

McGregor, W. 1990. *A Functional Grammar of Gooniyandi*. Amsterdam: John Benjamins.

McIlwain, H. & Bachelard, H.S. 1985. *Biochemistry and the Central Nervous System*, Edinburgh: Churchill Livingstone.

McLeod, S. & Bleile, K. 2003. 'Neurological and developmental foundations of speech acquisition'. Paper presented at American Speech Language Hearing Association Convention, Chicago, USA.

McMahon, A. & McMahon, R. 2005. *Language Classification by Numbers*. Oxford: Oxford University Press.

McNeill, D. 1966. 'Developmental psycholinguistics'. In F. Smith & G. Miller (eds), *The Genesis of Language: A Psycholinguistic Approach* (pp. 15–84). Cambridge, MA: MIT Press.

Meakins, F. 2008. 'Unraveling languages: Multilingualism and language contact in Kalkaringi'. In J. Simpson & G. Wigglesworth (eds), *Children's Language and Multilingualism: Indigenous Language Use at Home and School* (pp. 247–64). New

York: Continuum.

Meakins, F. 2013a. ‘Gurindji kriol’. In S. Michaelis, P. Maurer, M. Haspelmath & M. Huber (eds), *Atlas of Pidgin and Creole Structures. Vol. II: The Language Surveys*. Oxford: Oxford University Press.

Meakins, F. 2013b. ‘Mixed languages’. In P. Bakker & Y. Matras (eds), *Contact Languages: A Comprehensive Guide* (pp. 159–228). Berlin: De Gruyter Mouton.

Meakins, F. 2013c. ‘The phonological forms and perceived functions of janyarrp, the Gurindji “baby talk” register’. *Lingua*, **134**, 170–93.

Meakins, F., McConvell, P. Charola, E., McNair, N., McNair, H. & Campbell, L. with Wightman, G. 2013. *A Dictionary of Gurindji*. Batchelor: Batchelor Press.

Mechelli A, Crinion, J.T., Noppeney, U., O’Doherty, J., Ashburner, J., Frackowiak, R.S. & Price, C.J 2004. Neurolinguistics: Structural plasticity in the bilingual brain. *Nature*, **431**, 757.

Meyerhoff, M. 2016. ‘Borrowing from Bislama into Nkep (East Santo, Vanuatu): Quantitative and qualitative perspectives’. *Languages and Linguistics in Melanesia*, **34**, 77–94.

Mikolov, T., Chen, K., Corrado, G. & Dean, J. 2013. ‘Efficient estimation of word representations in vector space’. *arXiv,* https://arxiv.org/abs/1301.3781.

Milroy, J. 1992. *Linguistic Variation and Change*. Oxford: Blackwell.

Milroy, J. 1993. ‘On the social origins of language change’. In C. Jones (ed.), *Historical Linguistics: Problems and Perspectives* (pp. 215–36). New York: Longman.

Milroy, J. & Milroy, L. 1985. *Authority in Language: Investigating Language Prescription and Standardisation*. New York: Routledge and Kegan Paul.

Milroy, L. 1987. *Language and Social Networks*. New York: Blackwell.

Mistica, M., Baldwin, T., Cordella, M. & Musgrave, S. 2008. ‘Applying discourse analysis and data mining methods to spoken OSCE assessments’. In *Proceedings of the 22nd International Conference on Computational Linguistics*. Vol. 1 (pp. 577–84). Association for Computational Linguistics.

Möller, J., Jansma, B., Rodriguez-Fornells, A. & Münte, T.F. 2007. ‘What the brain does before the tongue slips’. *Cerebral Cortex*, **17**(5), 1173–8.

Moorfield, J.C. & Ka’ai, T.M. 2009. *Kupu Arotau – eweri tāima: Loanwords in Māori 1842–1952. Some Interesting Features of Loanwords Collected from the Māori Language Newspapers, the Paipera Tapu and Some early Published Writings.* Auckland: Te Ipukarea, the National Māori Language Research Institute, Auckland University of Technology, https://www.researchgate.net/publication/48668173_Nga_Kupu_Arotau_-_eweri_taima_Loanwords_in_Maori_1842-1952, Morey, S. 2010. *Turung: A Variety of Singpho Language Spoken in Assam*. Canberra: Pacific Linguistics.

Musgrave, S. 2012. ‘An introduction to the Australian National Corpus project’. *Humanities Today*, **3**, 29–34.

Musgrave, S. & Burridge, K. 2010. ‘[jɛəs prɔɪm mɪnɪstə]’. Paper presented to the Thirteenth Australasian International Conference on Speech Science and Technology 2010, Melbourne, 14–16 December.

Musgrave, S. & Burridge, K 2014. ‘Bastards and buggers: Historical snapshots of Australian English swearing patterns’. In K. Burridge & R. Benczes (eds), *Wrestling with Words and Meanings: Essays in Honour of Keith Allan* (pp. 3–32). Melbourne: Monash University Publishing.

Mushin, I. & Gardner, R. 2009. ‘Silence is talk: Conversational silence in Australian Aboriginal talk-in-interaction’. *Journal of Pragmatics* **41**(10), 2033–52.

Mushin, I. & Watts, J. 2016. ‘Identifying the grammars of Queensland ex-government Reserves: The case of Woorie Talk. In F. Meakins & C. O’Shannessy (eds), *Loss and Renewal: Australian Languages Since Colonisation* (pp. 57–85). Berlin: De Gruyter Mouton

Myers-Scotton, C. 2002. *Contact Linguistics: Bilingual Encounters and Grammatical Outcomes*. Oxford: Oxford University Press.

Myers-Scotton, C. & Jake. J.L. 2009. ‘Which language? Participation potentials across lexical categories in codeswitching’. In L. Isurin, D. Winford & K. de Bot (eds), *Interdisciplinary Approaches to Code-Switching* (pp. 206–40). Amsterdam: Walter Benjamins.

Nash, D. & Simpson, J. 1981. ‘No-name’ in central Australia’. In C.S. Masek, R.A. Hendrick & M.F. Miller (eds), *Proceedings on Language and Behavior* (pp. 165–77). Chicago: Chicago Linguistics Society.

Nayler, B.S. 1869. *Commonsense Observations on the Existence of Rules Regarding the English Language*. Melbourne: B.S. Nayler.

Nemoto, N. 1999. ‘Scrambling’. In N. Tsujimura (ed.), *The Handbook of Japanese Linguistics* (pp. 121–53). Oxford: Blackwell.

Nibun, Y. & Wigglesworth, G. 2014. ‘Early pragmatic differentiation in Japanese and German: A case study of a developing trilingual child in Australia’. *International Journal of Multilingualism*, **11**(1), 76–96.

Niedzielski, N.A. & Preston, D.R. 2000. *Folk Linguistics*. Berlin: Mouton de Gruyter.

Noonan, P. 1998. ‘Toward candor and courage in speech’. In K.A. Ackley (ed.), *Essays from Contemporary Culture* (pp. 368–76). Fort Worth, TX: Harcourt Brace and Co.

Norris, D., McQueen, J.M., Cutler, A., Butterfield, S. & Kearns, R. 2001. ‘Language-universal constraints on speech segmentation’. *Language and Cognitive Processes,* **16**, 637–60.

Oaks, D.D. 2001. *Linguistics at Work: A Reader of Applications*. Cambridge, MA: Heinle and Heinle.

Ochs, E. 1979. ‘Transcription as theory’. In E. Ochs & B. Schiefflen (eds), *Developmental Pragmatics*. New York: Academic Press.

Ochs, E. & Schieffelin, B. 2008. ‘Language socialization: A historical overview’. In P.A. Duff & N.H. Hornberger (eds), *Encyclopedia of Language Education*. 2nd edn. Vol. 8 (pp. 3–15). New York: Springer.

Ogiermann, E. 2009. ‘Politeness and in-directness across cultures: A comparison of English, German, Polish and Russian requests’. *Journal of Politeness Research,* **5**(2),189–242.

Olsson, J. 2004. *Forensic Linguistics: An Introduction to Language, Crime and the Law*. London: Continuum.

Otheguy, R., García, O. & Reid, W. 2015. ‘Clarifying translanguaging and deconstructing named languages: A perspective from linguistics’. *Applied Linguistics Review,* **6**(3), 281–307.

Pavlenko, A. & Blackledge, A. (eds) 2004. *Negotiation of Identities in Multilingual Contexts*. Clevedon: Multilingual Matters.

Pawley, A. 2008. Australian vernacular English: Some grammatical characteristics. In K. Burridge & B. Kortmann (eds), *Varieties of English: The Pacific and Australasia*. Berlin: Mouton de Gruyter, pp. 362–97.

Peal, E. & Lambert M. 1962. ‘The relation of bilingualism to intelligence’. *Psychological Monographs,* **75**, 1–23.

Pearce, W. & Stockings, E. 2011. ‘Oral narratives produced by Aboriginal Australian children: Dilemmas with normative comparisons’. *Acquiring Knowledge in Speech, Language and Hearing,* **13**(3), 126–31.

Pennycook, A. & Otsuji, E. 2015. *Metrolingualism Language in the City*. New York: Routledge.

Pereira, F., Lou, B. Pritchett, B. Ritter, S., Gershman, J., Kanwisher, N. Botvinick, M. & Fedorenko, E. 2018. ‘Toward a universal decoder of linguistic meaning from brain activation’. *Nature Communications,* **9**(1). doi: 10.1038/s41467-018-03068-4.

Pereira, L., Manguilimotan, E., & Matsumoto, Y. 2016. Leveraging a large learner corpus for automatic suggestion of collocations for learners of Japanese as a second language. *CALICO Journal,* **33**(3), 311–32.

Perrott, D. 2014. ‘Communication disorders and mental health: The link we can’t ignore'. In K. Burridge & R. Benczes (eds), *Wrestling with Words and Meanings* (pp. 262–82). Melbourne: Monash University Publishing.

Peters, A. 1983. *The Units of Language Acquisition*. Cambridge: Cambridge University Press.

Peters, P. 2004. *The Cambridge Guide to English Usage*. Cambridge: Cambridge University Press.

Peters, P. 2007. *The Cambridge Guide to Australian English Usage*. Melbourne: Cambridge University Press.

Peterson, D.J. 2015. *The Art of Language Invention: From Horse-Lords to Dark Elves, the Words Behind World-building*. New York: Penguin.

Pinker, S. 2007a. *The Language Instinct: How the Mind Creates Language*. New York: HarperCollins.

Pinker, S. 2007b. *The Stuff of Thought: Language as a Window into Human Nature*. New York: Penguin.

Plaza, M, Gatignol, P., Leroy, M. & Duffau, H. 2009. 'Speaking without Broca's area after tumor resection'. *Neurocase*, **15**(4), 294–310.

Puckett E. 2014. 'Japanese discourse particles', MA thesis, University of Oregon.

Pullum, G.K. 2005. 'The continuing misrepresentation of FOXP2 effects'. *Language Log*, 5 September, http://itre.cis.upenn.edu/~myl/languagelog/archives/002456.html

Radford, A. 2004. *English Syntax: An Introduction*. Cambridge: Cambridge University Press.

Radford, A. 2009. *Transformational Grammar: A First Course*. Cambridge: Cambridge University Press.

Rasinger, S.M. 2008. *Quantitative Research in Linguistics: An Introduction*. London: Continuum.

Reddy, M.J. 1979. 'The conduit metaphor: A case of frame conflict in our language about language'. In A. Ortony (ed.), *Metaphor and Thought* (pp. 284–310). Cambridge: Cambridge University Press.

Rezanova, Z.I. & Kogut, S.V. 2015. 'Types of discourse markers: Their ethnocultural diversity in scientific text'. *Procedia – Social and Behavioral Sciences*, **215**, 266–72.

Rickford, J.R. 1999. 'The ebonics controversy in my backyard: A sociolinguist's experiences and reflections'. *Journal of Sociolinguistics*, **3**(2), 267–75.

Ridley, A.M., Gabbert, F. & La Rooy, D.J. (eds) 2012. *Suggestibility in Legal Contexts: Psychological Research and Forensic Implications*. London: Wiley-

Blackwell.

Riehl, C.M. 2010. ‘The mental representation of bilingualism’, *WIREs Cognitive Science*, https://onlinelibrary.wiley.com/doi/abs/10.1002/wcs.74.

Robertson, W. 2017. ‘He’s more Katakana than Kanji: Indexing identity and self-presentation through script selection in Japanese manga (comics)’. *Journal of Sociolinguistics*, **21**(4), 497–520.

Rodriguez Louro, C. 2013. ‘Quotatives down under: Be like in cross-generational Australian English speech’. *English World-wide: A Journal of Varieties of English*, **34**(1), 48–76.

Rogers, H. 2004. *Writing Systems: A Linguistic Approach*. Oxford: Blackwell.

Ross, N.E. 1928. *How to write telegrams properly*. Girard, KS: Haldeman-Julius.

Salzman, M. & Halloran, M. J. 2004. ‘Culture, meaning, self-esteem and the re-construction of the cultural worldview’. In J. Greenberg, S.L. Koole & T. Pyszczynski (eds), *Handbook of Experimental Existential Psychology* (pp. 231–46). New York: Guilford Press.

Sapir, E. 1921. *Language: An Introduction to the Study of Speech*. New York: Harcourt, Brace and World.

Saville-Troike, M. 2003. *The Ethnography of Communication*. 3rd edn. Oxford: Blackwell.

Schneider, E. 2004. ‘Global synopsis: phonetic and phonological variation in English world-wide’, in E. Schneider, K. Burridge, B. Kortman, R. Mesthrie & C.Upton (eds), *A Handbook of Varieties of English*, Vol. **1** (pp. 1111–38). Berlin: Mouton de Gruyter.

Sebba, M. 2006. ‘Ideology and alphabets in the former USSR’. *Language Problems and Language Planning*, **30**(2), 99–125.

Sharifian, F. 2002. ‘Chaos in Aboriginal English discourse’. In A. Kirkpatrick (ed.), *Englishes in Asia: Communication, Identity, Power and Education* (pp. 125–41). Melbourne: Language Australia.

Sharifian, F., Rochecouste, J. & Malcolm, I.G. 2004. ‘“But it was all a bit confusing …”: Comprehending Aboriginal English texts’. *Language, Culture and Curriculum*, **17**(3), 203–28.

Sharples, P. 2013. ‘State of the Māori nation’. Speech at Maori of the Year ceremony, 7 February, http://www.beehive.govt.nz/speech/state-maori-nation.

Schiffrin, D. 1987. *Discourse Markers*. Cambridge: Cambridge University Press.

Severin, A.A. 2017. ‘The nature of prescriptivism and descriptivism online: The case of Reddit and r/badlinguistics’. Unpublished PhD thesis, Monash University.

Shulova-Piryatinsky, I. & Harkins, D.A. 2009. ‘Narrative discourse of native and immigrant Russian-speaking mother–child dyads’. *Narrative Inquiry*, **19**(2), 328–55.

Siegel, J. 2008. *The Emergence of Pidgin and Creole Languages*. Oxford: Oxford University Press.

Siegel, J. 2012. *Second Dialect Acquisition*. Cambridge: Cambridge University Press.

Sigley, R. & Holmes, J. 2002. ‘Looking at girls in corpora of English’. *Journal of English Linguistics*, **30**(2): 138–57.

Silby, W. 2008. ‘t-flapping in New Zealand English: *A change of time?’ New Zealand English Journal*, **22**, 24–35.

Simard, L.M.D., Taylor, M. & Giles, H. 1976. ‘Attribution processes and interpersonal accommodation in a bilingual setting’. *Language and Speech*, **19**(4), 374–87.

Simpson, J. 2008, ‘Language landscapes of children in remote Australia’. In R. Amery & J. Nash (eds), *Inaugural Indigenous Languages Conference (ILC)* (pp. 51–60). Adelaide: University of Adelaide.

Slobin, D.I. (ed.) 1985–97. *A Cross-Linguistic Study of Language Acquisition*. Hillsdale, NJ: Lawrence Erlbaum.

Smith, G.W. 1998. The political impact of name sounds. *Communications Monographs*, **65**(2), 154–72.

Smith, N.V. 2002. *Language, Banana Sand Bonobos: Linguistic Problems, Puzzles and Polemics*. Oxford: Blackwell.

Spencer, A. & Zwicky, A. (eds) 2001. *The Handbook of Morphology*. Oxford: Wiley-Blackwell.

Sperber, D. & Wilson, D. 1995. *Relevance: Communication and Cognition*. 2nd edn. Oxford: Blackwell.

Stebbins, T.N. 2004. ‘Mali Baining perspectives on language and culture stress’. *International Journal of the Sociology of Language (Small Languages and Small Language Communities)*, **169**, 161–75.

Stebbins, T.N. 2011. *Mali (Baining) Grammar*. Canberra: Pacific Linguistics.

Stebbins, T.N. & Planigale, M. 2010. ‘Explaining the unknowable: Accessibility of meaning and the exegesis of Mali Baining songs’. *Australian Journal of Linguistics*, **30**(1), 141–54.

Stebbins, T.N. with the assistance of Tayul, J. 2009. *Mali (Baining) Texts*. Canberra: Pacific Linguistics.

Stebbins, T.N. with the assistance of Tayul, J 2012. *Mali (Baining) – English Dictionary*. Canberra: Pacific Linguistics.

Stephens, R. 2013. 'Swearing: The language of life and death'. *The Psychologist*, **26**(9), 650–3.

Stern, G. 1965 [1931]. *Meaning and Change of Meaning (with Special Reference to the English Language)*. Bloomington, IN: Indiana University Press.

Stroop, J.R. 1935. 'Studies of interference in serial verbal reactions'. *Journal of Experimental Psychology*, **18**(6), 643–62.

Strunk, W. Jr & White, E.B. 1979. *The Elements of Style*. 3rd edn. New York: Macmillan.

Stubbs, M. 1995. 'Collocations and cultural connotations of common words'. *Linguistics and Education*, **7**: 379–90.

Sullivan, K. 2018 *Mixed Metaphors: Their Use and Abuse*. London: Bloomsbury.

Svartvik, J. 1968. *The Evans Statements: A Case for Forensic Linguistics*. Goteborg: University of Goteborg.

Swanson, L.W. 1995. 'Mapping the human brain: Past, present and future'. *Trends in Neurosciences*, **18**(11), 471–4.

Swinney, D. 1979. 'Lexical access during sentence comprehension: (Re)consideration of context effects'. *Journal of Verbal Learning and Verbal Behavior*, **18**, 645–59.

Szakay, A. 2008. *Ethnic Dialect Identification in New Zealand: The Role of Prosodic Cues*. Saarbrucken: VDM Verlag.

Tagliamonte, S. 2012. *Variationist Sociolinguistics Change, Observation, Interpretation*. Chichester: Wiley-Blackwell.

Takeuchi, M. 2006. *Raising Children Bilingually through the 'One Parent-One Language' Approach: A Case Study of Japanese Mothers in the Australian Context*. Bern: Peter Lang.

Tallerman, M. 2015. *Understanding Syntax*. 4th edn. New York: Routledge.

Tan, L.H., Chan, A.H.D., Khong, P.L., Yip, L.K.C. & Luke, K.K. 2008. 'Language affects patterns of brain activation associated with perceptual decision'. *Proceedings of the National Academy of Sciences*, **105**(10), 4004–9.

Tannen, D. 1984. *Conversational style: Analyzing talk among friends*. Norwood, NJ: Ablex.

Taylor, T.J. 1990. 'Which is to be master? The institutionalization of authority in the science of language'. In J.E. Joseph & T.J. Taylor (eds), *Ideologies of Language* (pp. 9–26). London: Routledge.

Tayul, J., Stebbins, T.N. & Marunga Mali Language Project (2004). *Asek dēchama mambu nēva bang – Stories and Songs from the Village.* Melbourne: Planigale Publications, in conjunction with Marunga Mali Language Project.

Thomason, S. 2013. ‘Ultraconserved words? Really??’ *Language Log*, 8 May, http://languagelog.ldc.upenn.edu/nll/?p=4612.

Thurlow, C. 2003. ‘Generation txt? The sociolinguistics of young people’s text-messaging’, http://extra.shu.ac.uk/daol/articles/v1/n1/a3/thurlow2002003-paper.html.

Tognini-Bonelli, E. 2001. *Corpus Linguistics at Work*. Amsterdam: John Benjamins.

Tokuhama-Espinosa, T. 2003. *The multilingual mind. Westport*, Westport, CT: Praeger.

Traugott, E.C. & Dasher, R.B. 2002. *Regularity in Semantic Change*. Cambridge: Cambridge University Press.

Trench, R.C. 1857. *On Some Deficiencies in Our English Dictionaries. Papers Delivered to the Philological Society’s November 1857 Meeting*; http://public.oed.com/history-of-the-oed/archived-documents.

Trudgill, P. 1975. *Accent, Dialect and the School*. London: Edward Arnold.

Trudgill, P. 2004. *New-Dialect Formation: The Inevitability of Colonial Englishes*. Oxford: Oxford University Press.

Turing, A.M. 1950. ‘Computing machinery and intelligence’. *Mind*, **59**(236), 433–60.

Twain, M. 1880. *The Awful German Language* (published as an appendix in the second volume of Twain’s *A Tramp Abroad*). Toronto: Robertson.

United Kingdom Department of Education and Science 1988. *Report of the Committee of Inquiry into the Teaching of English Language.* London: HMSO.

United Nations 1948. *Universal Declaration of Human Rights*, http://www.refworld.org/docid/3ae6b3712c.html.

Urban, H.B. 1998. ‘The torment of secrecy: Ethical and epistemological problems in the study of esoteric traditions’. *History of Religions*, **37**(3), 209–48.

von Frisch, K. 1967. *The Dance Language and Orientation of Bees*. Cambridge, MA: Harvard University Press.

Wardhaugh, R. 2010. *An Introduction to Sociolinguistics*. 6th edn. Oxford: Wiley-Blackwell.

Warren, P. 2013. *Introducing Psycholinguistics*. Cambridge: Cambridge University Press.

Wells, J.C. 2000. ‘Overcoming phonetic interference’. *English Phonetics, Journal of the English Phonetic Society of Japan*, **3**, 9–21.

Wierzbicka, A. 1991. *Cross-Cultural Pragmatics: The Semantics of Human Interaction*. Berlin: Mouton de Gruyter.

Wierzbicka, A. 1992. *Semantics, Culture, and Cognition: Universal Human Concepts in Culture-Specific Configurations*. New York: Oxford University Press.

Willoughby, L. 2011. ‘Sign language users’ education and employment levels: Keeping pace with changes in the general Australian population?’ *Journal of Deaf Studies and Deaf Education*, **16**(3), 401–13.

Willoughby, L. 2012. ‘Language maintenance and the deaf child’. *Journal of Multilingualism and Multicultural Development*, **33**(6), 605–18.

Willoughby, L. 2014. ‘The role of professional advice in shaping language choice in migrant background families with deaf children’. In J. Hajek & Y. Slaughter (eds), *Challenging the Monolingual Mindset* (pp. 149–61). Clevedon: Multilingual Matters.

Wilson, A. 2010. ‘Gillard’s intellect and Blanchett’s voice?’ *Crikey*, 1 July, http://blogs.crikey.com.au/fullysic/2010/07/01/gillards-intellect-and-blanchetts-voice.

Winawer, J., Witthoft, N., Frank, M.C., Wu, L., Wade, A.R. & Boroditsky, L. 2007. ‘Russian blues reveal effects of language on color discrimination’. *Proceedings of the National Academy of Sciences*, **104**, 7780–5.

Winford, D. 2003. *An Introduction to Contact Linguistics*. Oxford: Blackwell.

Woolard, K.A & Schieffelin, B.B. 1994 ‘Language ideology’. *Annual Reviews in Anthropology*, **23**, 55–82.

Wray, A. & Bloomer, A. 2013. *Projects in Linguistics and Language Studies*. 3rd edn. London: Routledge.

Wren, M. 2009. ‘Experiences and attitudes concerning English spelling’. Unpublished BA(Hons) thesis, Monash University.

Zachrisson, R.E. 1931. ‘Four hundred years of English spelling reform’. *Studia Neophilologica*, **4**, 1–69.

Zion, L. 2007. *The Sounds of Aus*, one-hour documentary screened on ABC TV.

Zuckermann, G. 2013. Historical and moral arguments for language reclamation. *History and Philosophy of the Language Sciences*; https://hiphilangsci.net/2013/06/26/historical-and-moral-arguments-for-language-reclamation.

Index

a (allomorph), 126
Abbott, Tony, 169
Aboriginal Areas Protection Authority, 45
Aboriginal English, 149
  and Australian English, 307
  ethnic-related variation, 316
Aboriginal languages
  community language centres, 45–46
  disappearing, 390–391
  number of, 390
  revival, 499–500
  voicing, 181
academic literacy, 567
accents, 8, 181, 303, 485
acceptance, 504
accommodation theory, 401–403
acoustic analysis, 206, 209
acoustic phonetics, 206–213
acquired disability, 542–543
acronyms, 94–95
Acton, Peter, 50
Adams, Douglas, 19, 406
Adams, Robert, 404
addition of sounds, 223–224
-*ade* (ending), 140
adjacency pairs, 436
adjective phrases, 278
adjectives, 135
  haplology, 223
  to nouns, 101, 133
  to verbs, 101, 134
adverbials, 260, 284, 287–288
adverbs, 136
advertising, 46–47, 153
affixation, 98–99, 116, 133, 137
affixes, 116, 118–119
  bound, 118
  types of functions, 120–126
affricates, 183
African American English, 71–72, 473
age-related differences, 312, 313–316
ageing, 65

agglutinating languages, 27, 347
agrammatism, 542
agreement, 135
agreement maxim, 466
airline industry, 50
Algeo, John, 112
algorithms, 587, 598
Allen, Wendy, 472
alliteration, 554
allomorphs, 126–128, 243–245
allomorphy, 589
allophones, 217–222
    complementary distribution, 231–237
alphabet, 252
alphabetic systems, 11, 249
alphabetisms, 94
alveolar sounds
    consonants, 180, 183
    stops, 182
ambiguities, 50, 149–150, 454
    lexical, 535
    syntactic, 276–277
American Dialect Society, 102, 112
American English
    in Australian English, 486–490
    r sound, 224, 344
an (allomorph), 126
Anabaptists, 381, 395, 471, 482
analogy, 161, 349–351, 559–560
anaphora, 424, 425
angle brackets, 196, 217
animal communication, 5
    indexical signs, 8
    teaching human communication, 15
anomaly, 150
anomia, 542
Anomic aphasia, 132
anonymity, 74–75
AntConc (software), 593–594
anticipatory assimilation, 226
anti-languages, 326
antonomy, 151
aphasia, 131–132, 542–543
apostrophes, 133
approbation maxim, 466
approximants, 184–186
apps, 582, 584
apraxia, 542
Arabic, 138, 144
archaic words, 108
areal linguistics, 386–387

Aristotle, 161, 240
arrows,
  use of, 235, 277, 314, 337
articles, 129
articulation
  manner of, 182
  places of, 179–181
articulators, 178
artificial languages, 51–52
ash (ending), 171
aspect, 124, 125
  in Russian, 30
aspiration, 196, 217
assimilation, 226–229, 340–343, 483, 506, 507
asterisks,
  use of, 337
attitudes, 370–372, 480
Aurora Project, 45
Auslan
  age-related differences, 312
  cat, 9
  cohesion, 425
  compounds, 97
  described, 21–22
  dog, 9
  and gestures, 475
  interpreters, 405–406
  regional variation, 311–312
  semantic roles, 295
  topic versus comment, 416
Austin, J. L., 456
Australian English
  and Aboriginal English, 307
  adhering to standard, 66–67
  Americanisation of, 486–490
  diphthongs, 193
  high rising tone or terminal (HRT), 314–315
  in Melbourne, 213
  regional variation, 309–311
  and standardisation, 504
  syntax, 73–74
  vowels, 189, 190–192, 211, 212–213
Australian Institute of Aboriginal and Torres Strait Islander Studies (AIATSIS), 25, 45
Australian National Corpus (AusNC), 53, 58, 61
Australian Questioning Intonation (AQI), 200
authenticity, 486
authors, 43
authorship, 74–75, 272–273
autism spectrum disorders (ASD), 541–542
Ayçiçegi-Dinn, A., 538

babble, 554
Babel Fish, 406
babies, 551, 554, 556, 564 *See also* toddlers
baby talk, 551, 552–553, 578
back-channelling, 435
backformation, 99–100
backronyms, 95
back-translation, 408
Baining ethnic group, 505 *See also* Mali
Baining language family, 27 *See also* Mali
Balipathar, 377
bank, 153
Banter Principle, 472–473
Barker, Ronnie, 203
Barnes, William, 490
basic clauses, 258–261, 285
basic literacy, 567
bathtub effect, 532
Bauer, Laurie, 96, 121, 136
be (verb), 125–126, 139
bees, 5, 8
behavioural methods, 525–527
benefits, 72
Berko Gleason, Jean, 556, 559, 560
berry fruit, 119–120
bias of language structure, 261
Biber, Douglas, 597
Big Australian Speech Corpus, 53, 58
Big Brother (television show), 73–74
bilabial sounds
    consonants, 180, 182
    stops, 182
bilingualism, 382, 573, 574–577
    and brain, 537–539
    code-switching, 384–385
Blake, Barry, 153, 548
blends, 100–101
bloody, 60, 288–289
Bolinger, Dwight, 37
Bontoc, 145
borrowings, 104–107, 142, 381
    structural, 385–386
bottom–up models, 528
Bowern, Claire, 599
boy, 61, 77
brackets
    and constituency, 267, 268, 276
    use of, 196, 217, 232, 278, 280
brain
    and bilingualism, 537–539
    damage, 543, 553

  development, 550, 553
  and language, 520–524
  structure, 520–522
brain maps, 543
brand names, 39, 47, 101, 206, 221
British English
  accents, 181, 485
  in Australia, 309, 504
British National Corpus, 592, 593, 594
bro, 96
broadening, 164
Broca’s aphasia, 131, 543
Broca’s area, 520, 521, 522
brother, 96
Brown, P., 467, 470
Brown, Roger, 556, 558
Brown Corpus, 592
bugger, 59
Burke, Isabelle, 73
Burnham, Denis, 53
Burridge, Kate, 23, 62, 471, 486, 500
business
  linguists in, 49–51
Butcher, Andrew, 307
Butler, Judith, 508
Bybee, Joan, 363

Cai, Peter, 49
Caldwell-Harris, C. L., 538
calques, 105
Calude, Andrea, 595
Cameron, Deborah, 75, 486, 493, 508
careers
  in linguistics, 37–53
  in phonetics, 204–205, 209–210
and phonology, 220
carving (lexicon of), 159
case, 123, 274, 275
  German, 23
case studies, 63
casual style, 320–321
cat (Auslan sign), 9
categories, 270–273
category labels, 270
-ceive (ending), 118
Celtic, 388
cerebrum, 520–522
changes in language
  across time, 20–21, 336–337
  attitudes, 370–372, 480
  contact-induced, 385–386, 388–389

  gradual, 358–359
  grammar, 347–355, 359, 363
  how, 355–359
  mapping, 369–370
  phonology, 337–346, 359
  terminology, 327
  variation, 302
  what, 337–355
  why, 359–363
chatbots, 600
Chaucer, Geoffrey, 172
Chen, Keith, 518
child development, 550
  reading, 567–569, 570, 577
  writing, 569
child developmental disabilities, 540–542
child directed speech (CDS), 551–553
child language acquisition
  ages and stages, 554–556, 558, 564, 572
  analytical pathway, 550
  approaches, 550–551
  bilingual, 574–577
  critical period hypothesis, 553–554
  differentiation, 575–577
  fis phenomenon, 556
  grammar, 558–560, 575, 577
  growing up bilingual, 574–577
  input, 551–553
  instructions, 563–564
  late emergence, 551
  phonology, 554–556, 575
  research approaches, 549
  vocabulary, 556–558, 575
children
  socially isolated without language, 554
  and spoken language, 529–530
      See also toddlers
Chinese
  dyslexic, 570
  influence on Japanese, 27
Chomsky, Noam, 18, 278
Chukchi, 86
circumfixes, 117
classes. *See* word classes
clauses
  basic, 258–261
  types of, 282–283
cleft constructions, 421–422
clicks
  consonants, 187
clipped words, 95–96

clitics, 120
Clyne, Michael, 378, 392
coda, 198
code-switching, 384–385
codification, 503
cognates, 364–365
Cohen, Bob, 46
coherence, 428
cohesion, 424–428
coinages, 93–107
collocation, 158–160
collocation analysis, 597
collocation research, 88
collocations, 595
colloquialisation, 326–327
Collyer, Eugenie, 45–46
colours, 519
comment, 416–417
commonisation, 101–102
communication
    competence, 18, 41
    effective, 44
    ethnography, 442–443
    gestures in, 473–476
    hybrid forms of, 12
    linguists in, 42–44
    types of, 431–437
communication disorders, 47–49, 220
communities
    and field linguists, 44–46
        See also speech communities
community language centres
    linguists in, 45–46
comparative method, 364
compensatory lengthening, 337
competence, 18, 41
complement clause, 264
complementary antonyms, 151
complementary distribution, 126, 231–237
complements, 260, 283
completed (aspect), 125
complex intransitive clauses, 283
complex sentences, 261, 265, 563
complex transitive clauses, 283
componential analysis, 155
compounds, 96–98
    reduplication, 103–104
comprehension, 533–535
    and sentences, 536
computational linguistics, 52
computing

  - linguists in, 52–53
- concordances, 593
- conduit metaphor, 451–452
- conjunction, 425
- conlangers, 52
- Connolly, Cyril Vernon, 449
- connotation, 148
  - changes, 164, 166
- consent, 61, 72, 73, 75
- consonants
  - affricates, 183
  - approximants, 184–186
  - assimilation, 243–244
  - and child language acquisition, 554
  - clicks, 187
  - deletion, 337–338
  - deletion of sounds, 222
  - devoicing, 342
  - fricatives, 183
  - International Phonetic Alphabet (IPA), 186–187
  - length of, 248
  - manner of articulation, 182
  - modifications, 138
  - places of articulation, 179–181
  - and spelling, 195
  - stops, 182–183
  - syllabic, 199
  - voicing, 181
- constituency
  - and brackets, 267, 268, 276
  - tests, 267–269
- constituents, 266–269, 270, 271
- constructed language creators, 52
- consultative style, 320–321
- contact between languages
  - types of, 377–381
- contact-induced changes, 385–386, 388–389
- contact languages, 397
- context
  - linguistic, 452–456
  - and meaning, 157–160
  - situational, 452–456
  - social, 19–20, 355–358
- continuants, 183
- continuous (tense), 125
- contour systems, 201
- contranyms, 154, 165
- contrastive distribution, 230
- conventionalisation, 8
- conventionalised gestures, 473
- convergence, 386, 402–403

conversational partners, 552
conversations, 433–437
    expressing support, 435
    recording, 73
    turn-taking, 436–437
conversions, 101, 133
Cook, Michael, 150
Cooperative Principle, 457–461, 464
coordinate sentences, 261, 262, 265
coordination, 259, 262
coordinators, 129
Corbett, Ronnie, 203
Cormann, Mathias, 453
corpora, 53, 58–61
corpus linguistics, 592–598
Corpus of Oz Early English (COOEE), 58
co-speech gestures, 11
Coupland, Nikolas, 402
coverbs, 143
covert prestige, 485
creation of grammar, 353–355
creation of languages, 52, 137, 381, 395–401
creation of words, 82, 93–107, 326
creative arts
    linguists in, 51–52
creoles, 396–400 *See also* Kriol
critical period hypothesis, 553–554
cross-linguistic influences, 385–386
Crowley, Terry, 240
Crystal, David, 104, 335
cultural differences, 438–441, 469–473
cuneiform, 250
curly brackets, 278, 280
Cushing, Steven, 50

Dahl, Roald, 118
dangling modifiers, 277
dashes,
    use of, 235
data
    experimental, 67–68
    interview, 65
    observational, 61–63
    sources, 58–68
    uses of, 69–71
de Saussure, Ferdinand, 8
Deaf people, 10, 42
declarative (clauses), 258–259, 450
definite articles, 261
deictic gestures, 473
deixis, 427–428

Delbridge, Arthur, 192
deletion of sounds, 337–338
demonstratives, 129
denotation, 147
denotatum, 147
dense networks, 313, 360
dental sounds
  consonants, 180, 183
dependents, 270
derivation, 120–126, 134, 137
derogatory expressions, 148–149
descriptive approaches, 15–17, 496
detached participle, 277
deterioration, 166
determiners, 129
Deutscher, Guy, 518
developmental disabilities, 540–542
diachronic linguistics, 336
diachronic perspective, 20, 355, 356
diacritic, 228
dialect chains, 308
dialects, 303, 304
  and languages, 306–308
  and misunderstandings, 305
  non-standard, 496, 497–498
  regional, 308–309
dichotic listening, 522
dictionaries, 16–17
  and compounds, 97
  and Google, 102
  of loanwords, 107
and meanings, 146
diglossia, 381–383
diphthongs, 193–195
direct objects, 286
directness, 465–466
dirty words. *See* taboo words
disambiguate, 149
discourse analysis, 41, 428–430, 596–598
discourse markers, 315, 429–430
discourse strategies, 412–418
  special, 418–424
discourse styles, 41, 48
dislocation, 423–424
displacement, 5
dissimilation, 228–229, 343
distinctive features, 241–243
ditransitive clauses, 283
divergence, 403
Diwurruwurru-jaru Aboriginal Corporation (the Katherine Regional Language Centre), 25

doctor–patient communication, 48
dog, 9, 124
dom (ending), 140
Dorian, Nancy, 492
Dothraki language, 52
Douglas, Mary, 483
Douglas, Susan, 542
Down Syndrome, 554
duality of patterning, 4
duals
  in Mali, 133
Dutch, 308, 364
dyslexia, 570

Eades, Diana, 37, 150, 292, 305, 405, 455
Eagleson, Robert, 272
Eastman, George, 93
Easton, Catherine, 500
echo question, 267
e-communication, 432
economic change, 393
editors, 43
education
  and language shift, 394
  linguists in, 40–42
  and non-standard variation, 497–498
eggcorn, 170
Einstein, Albert, 550
elaboration, 503
electrodermal monitoring, 328, 527
electro-encephalography (EEG), 527
electronic exchanges, 328–330
elevation, 166–167
Elliot, T. S., 146
ellipsis, 424
Elvish language family, 51–52
emergentist, 549
emojis, 475–476
emoticons, 12
emotions and words, 328, 441, 538–539
endangerment, 44
end-focus,
  principle of, 415
end-weight,
  principle of, 415
English
  borrowings from French, 105–106
  borrowings from German, 104–105
  borrowings from Russian, 30
  borrowings from Scandinavian, 106
  borrowings into German, 105

  borrowings into Japanese, 27, 107
  borrowings into Mali, 142
  changes in, 19
  described, 22–23
  historical changes, 336–337
  as international language, 443
  variation, 300–303
    See also history of English
Environment Canada, 591
epenthesis, 223, 338–339
eponyms, 102
er (ending), 140
error analysis, 530–532
espoused models, 62
ethics, 61, 71–75
ethnic groups, 505
ethnic-related variation, 316–317
ethnography of communication, 442–443
ethnolects, 316
etymology, 154
  bogus, 171
euphemisms, 163, 168, 323
evaluative models, 62
Evans, Nick, 390
existentials, 420
expansion, 552
experimental data, 67–68
expletives, 118, 327
expressions
  derogatory, 148–149
  formulaic, 87–89
  institutionalised, 88
  non-referring, 147
  referring, 147
extinction, 390–391
extraposition, 421
eye tracking, 526

f sound, 181
face, 467
face threatening acts, 467
Facebook chat, 12, 468
facework, 467
family trees, 369
Faroese, 225
Fawlty Towers (TV series), 320
Feagin, Crawford, 64
female speech, 495
  high rising tone or terminal (HRT), 200, 315
  vocal fry, 202
Ferguson, Naomi, 488

field linguists, 44–46, 458
Filipino nurses, 429
First Germanic Consonant Shift, 366
Firth, J. R., 158
fis phenomenon, 556
flapping, 218, 357–358
Florey, Margaret, 62
fneeze, 172
focus groups, 65
folk etymology, 154, 170
food
    borrowings from French, 106
force of words, 456–457
foreign language learning, 219, 574
foreign language teaching, 40–42 *See also* bilingualism
forensic linguistics, 39, 221–222, 322, 329–330
forensic phonetician, 204–205
forensic phonetics, 204–205, 209–210
Forensic Phonetics Australia, 205
formal style, 320–321
formality,
    levels of, 320–321
formants
    acoustic analysis, 211–213
    frequencies, 206–209
formulaic expressions, 87–89
Fraser, Helen, 204, 534
free variation, 237–239
French
    borrowings into English, 105–106
    changes in, 20
    deletion of sounds, 222
    vowels, 189
French, Peter, 210
frequency factors, 362–363
fricatives, 183, 224
Frisian, 306
Frommer, Paul, 52
front focus,
    principle of, 417–418
fronting, 268, 422–423
frozen style, 320–321
Fry, Stephen, 5, 87, 150
function labels, 270
functional literacy, 567
functional magnetic resonance imaging (fMRI), 527
functions, 270–273

g sound, 345–346, 356–357
Gage, Phineas, 523
galvanic skin responses, 527

garden path sentences, 536
Garner, James Finn, 414
gating techniques, 526
Gawne, Lauren, 476
gay, 163
geminate, 226
gender classes, 123, 124
gender differences, 319, 494–495, 507–508 *See also* female speech
generative grammar, 278
generativist approaches, 549
generosity maxim, 466
genetic affinity, 363
genetic relationships, 363–370
genetics research, 598–599
German
  acronyms, 94
  affixation, 98
  allomorphs, 128
  backformation, 100
  blends, 101
  borrowings from English, 105
  borrowings into English, 104–105
  case, 23, 275
  circumfixes, 117
  cognates, 364, 365
  collocation, 159
  compounds, 96
  consonants, 342
  conversions, 101
  described, 23
  dialects, 23
  diphthongs, 195
  and Dutch, 308
  epenthesis, 338
  eponyms, 102
  grammar, 275–276
  long words, 110
  metathesis, 344
  nasal sounds, 337–338
  nouns, 94
  phonemes, 236–237
  reduplication, 103
  shortenings, 95
  spelling reforms, 502
  suffixes, 117
  umlauting, 228, 342
  vowels, 247, 341
  word order, 290
German speakers
  learning English, 221
gerontolescence, 65

gestalt approach, 551
gestures, 11, 362
    in communication, 473–476
    and sign language, 474–475
Giles, Howard, 403
Gillard, Julia, 225, 495
girl, 61, 77, 147
given information, 413–415
Global Web-Based English, 592
glottals sounds
    consonants, 180
    stops, 182
Goffman, Erving, 467
gonna, 353–354
google (verb), 102
Google Books corpus, 592
Google Translate, 592
Gooniyandi, 291–292
goosebumps, 527
Gordon-Read, Annette, 49
gr words, 10
gradable antonyms, 151
Grainger, Percy, 490
grammar, 257
    acquisition of, 292, 558–560, 575, 577
    changes in, 347–355, 359, 363
    creation, 353–355
    defined, 491
    and styles, 323
        See also syntax
grammatical properties, 85
grammatical relations, 273–277
grammatical sentences, 14, 258
grammatical words, 87, 128
grammaticalisation, 353–355
grammaticality, 257–258
Grant, Tim, 330
graphisation, 501–502
Gresham's Law, 168
Grice, Paul, 459, 463, 464
Grimm's Law, 366
Gurindji
    baby talk, 552–553
    complex sentences, 295
    coverbs, 143
    described, 24–25
    diphthongs, 195
    kinship terms, 515, 545
    names, 90
    pronouns, 466
    reflexive pronouns, 426

  - use of, 379
  - variation, 307
  - vowels, 193
- Gurindji Kriol, 25, 307, 401, 407

- HAL (computer), 582
- Hall, Rich, 111
- Halloran, Michael, 392
- Hanson, Pauline, 162
- haplology, 223, 338
- Harrington, Jonathan, 210
- HCE system of transcription, 191, 192
- head (nouns), 270
- health sciences
  - linguists in, 47–49
- Heath, Chip, 47
- Heath, Dan, 47
- *Helsinki Corpus of English Texts*, 58
- hemispheres (brain), 521, 522
- Hendery, Rachel, 315
- hereby test, 457
- hesitations, 530
- high rising tone or terminal (HRT), 200, 314–315
- historical linguistics, 336
- history of English, 20–21, 397–398
  - resources, 58
- Hockett, Charles, 4
  - design features of language, 4–6
- *Hollowmen* (TV series), 265
- Holmes, Janet, 61, 77, 357, 494
- homographs, 152
- homonymy, 152–153, 154, 212
- homophones, 152, 219
- homorganic, 226
- honorifics, 27, 566–567
- *house that Jack built*, 281
- Householder, Fred, 449
- Hudson, Richard, 44, 292
- Hughes, Geoffrey, 108
- human–computer interfaces, 52
- human languages, 5
- human research ethics, 72–74
- humanities education
  - and business, 49–51
- Humboldt, Wilhelm von, 516
- humour, 152 *See also* jokes
- Hymes, Dell, 442
- hyperbole, 166, 167, 462
- hyphens, 97
- hyponymy, 155

iconic gestures, 473
iconicity, 7
identity functions of language, 483–485
ideologies of language, 479, 504, 507–508
ideophones, 7, 33
idiolects, 329
idioms, 88, 160
illocutionary acts, 456
imperative, 450
implicature, 464–465
impoliteness, 465–468
impositions
    ranking, 468
inalienable possessions, 89
indefinite articles, 261
indirect objects, 286
indirect speech acts, 457
indirectness, 470
inferencing, 464
infixes, 118
inflection, 120–126
    paradigms, 122–126
    plural markings, 130–131
inflectional languages, 348
informalisation, 326–327
information packaging, 412–418
informed consent, 61, 72, 73, 75
initialisms, 94
input, 549, 551–553
institutionalised expressions, 88
interactive, 528
interactive gestures, 473
interference, 572
    phonological, 220
    pragmatic, 573–574
interjections, 185–186, 246
International Phonetic Alphabet (IPA)
    consonants, 186–187
    vowels, 188
internationalism, 506
interpreters, 405–406
interpreting, 403, 405
interrogatives, 291–292, 450
interruptions, 446, 471
interview data, 65
interviews, 64–67
intimate style, 320–321
intonation, 197, 200, 554
intransitive clauses, 282
inversion, 422
Irish English, 16

irony, 454, 462
isogloss, 369
isolating languages, 347, 349
it (dummy pronoun), 421

James, Henry, 34
Japanese
  agglutinating, 347
  and assimilation, 483
  borrowings from English, 27, 107
  broadening, 164
  case studies, 63
  Chinese influence, 27
  consonants length, 248
  described, 26–27
  deterioration, 166
  discourse markers, 430
  dissimilation, 343
  grammatical relations, 293
  honorifics, 324, 566–567
  ideophones, 33
  kinship terms, 515
  learning, 88
  loanwords, 107, 235
  metathesis, 344
  narrowing, 165
  phonemes, 233–235
  phrase structure rules, 279
  and politeness, 469
  prefixes, 116–117
  reflexive pronouns, 426
  scrips, 249
  sentences, 291
  shifts, 165
  styles, 324
  topic versus comment, 417
  writing systems, 26
Japanese speakers
  learning English, 221
jargon, 319
Johnson, Samuel
  Dictionary of the English Language, 34
Johnston, T., 98, 311, 313, 425
jokes, 378, 425, 427, 460–461
Jones, Caroline, 552
Jones, William, 366
Joos, Martin, 320
Joyce, James, 93
justice, 72

Kamhmu

- morphological processes, 144

Kanno, Yasuko, 483
Katamba, Francis, 119
Kath and Kim (television series), 152
Kayardild, 391
Key Word in Context tools, 593
keyword analysis, 596
kinship terms, 156, 515–516, 545
Klingon
- language, 52
- morphology, 122

Koch, Harold, 150
Kodak, 93
Korean, 221
Kortmann, Bernd, 71
Koteyko, N., 159
Kriol, 253, 398–399, 401 *See also* Gurindji Kriol
Kruger, Paula, 53
Kubota, Ryuko, 443

l sound, 231–233
labiodental sounds
- consonants, 180, 183

labiovelar sounds, 184
Labov, William, 39–40, 64, 209, 330, 331, 357, 437
lag assimilation, 227
language
- and brain, 520–524
- classifying in use, 432
- design features, 4–6
- function of, 300
- origin of, 3
- survival, 492
- and thought, 517–519
- transmission of, 6

language acquisition, 548, 575
- delayed, 541
- pathways, 550–554
- research approaches, 549–550
  - *See also* child language acquisition; second-language acquisition (SLA)

language awareness, 43
language choices, 401–403
Language Creation Society, 52
language death, 387
language disabilities, 539–543
language families, 363–370
language maintenance, 380, 381–387
- and reclamation, 389–392

language modalities, 10–12
language planning, 499–504, 506–507

language processing
  models, 528–530
language production, 530–533
language shift, 380, 387–395
language skills, 43
language-switching, 384–385
Language in the Workplace Project, 50
languages
  agglutinating, 347
  artificial, 51–52
  case studies, 3, 21–32
  contact between, 377–381
  and dialects, 306–308
  differences, 515–517
  inflectional, 348
  isolating, 347, 349
  mixed, 400
  and nations, 505–507
  number of, 379
  polysynthetic, 86, 348
  researching variation, 331–332
  revival of, 392, 499–500
  secret, 245–246
  tone, 201–202, 248
and variation, 306
lateralisation, 521, 522, 524, 553
laterals, 184, 231
Latin, 366
law of semantic change, 168
lax vowels, 190
learning English, 221 *See also* language acquisition
learning foreign languages, 219
  strategies, 574
Leech, Geoffrey, 466
Leetspeak, 324
left-dislocation, 423
legal system
  and authorship, 272–273
  bias of language structure, 261
  and context, 455–456
  and forensic phonetics, 204–205, 209–210
  linguists in, 38–40, 92–93, 149
  online identities, 329–330
  styles, 321
length of sounds, 247
Levinson, S., 467, 470
Lewis, C. S., 166
lexemes, 115
lexical addition, 93–107
lexical cohesion, 424
lexical diffusion, 358, 362

lexical insertion rules, 279, 280
lexical markers, 318
lexical morphemes, 128–132
lexical relations, 150–155
lexical retrieval, 532
lexical signs, 475
lexical words, 86–87, 128
lexically conditioned, 128
lexicographers, 43
lexicography, 17
Lexicon (firm), 47
lexicons
  ins and outs of, 82
  obsolescence, 108
leximania, 82, 111
Liberman, Mark, 170
like, 315, 354
lingua franca, 379
linguistic awareness, 537
linguistic conditioning, 219
linguistic data
  sources, 58–68
  uses of, 69–71
linguistic determinism, 517
linguistic relativity, 517
linguistic variables, 314, 330

linguistics
  careers in, 37–53
  and Deaf people, 42
  forensic, 38, 221–222, 322, 329–330
and mental health, 48
linguists
  in advertising, 46–47
  in business, 49–51
  in communication, 42–44
  in computing, 52–53
  creative arts, 51–52
  in education, 40–42
  in health sciences, 47–49
  historical, 336
  in legal system, 38–40, 92–93, 149
  in marketing, 46–47
  in professional services, 49–51
  in technology, 52–53
  in the field, 44–46, 458
lips, 188, 189
liquids, 224, 226, 228, 554
lisps, 554
literacy,
  development of, 567–570
literal language rules, 455
Loakes, Debbie, 210, 212
loanwords, 247
  Japanese, 107, 235
    See also borrowings
locutionary acts, 456
Loftus, Elizabeth, 261
logical approaches, 549
logographic symbols, 250
logographic systems, 11, 249
Lucy ('Peanuts' character), 162
Lyman's Law, 343

machine translation, 591–592
MacKay, D. G., 526
Mackay, Sandra, 443
magneto-encephalography (MEG), 528
Maitz, Péter, 400
malapropisms, 169
Mali
  adjectives, 135
  adverbs, 137
  baby talk, 578
  basic clauses, 260
  borrowings from English, 142
  clauses, 290
  complement clause, 265

  - coordination, 262
  - derivation, 121
  - described, 27–28
  - inflection, 134
  - kinship terms, 516
  - Maxim of Quality, 459
  - Maxim of Quantity, 458
  - minimal pairs, 230
  - morphemes, 114–115, 129
  - morphological processes, 144
  - names, 90
  - narratives, 438, 440
  - noun class marking, 142
  - number marking, 133
  - offering tea, 63
  - order of words, 13
  - rate of syllables per second, 481
  - reflexive pronouns, 426
  - relative clauses, 264
  - and Tok Pisin, 382–383
  - topic versus comment, 416
  - transitivity, 294
  - verbs, 27
  - vowel harmony, 228
  - vowels, 193
  - word class alternations, 134
  - words, 83
- Mandarin Chinese
  - tone, 201
- manner. *See* Maxim of Manner
- manner diagrams, 182, 183, 184
- manner of articulation
  - consonants, 182
- Māori
  - compounds, 393
  - described, 28–29
  - educational initiatives, 45
  - kinship terms, 516
  - loanwords from English, 247
  - morphological processes, 145
  - multifunctional words, 136
  - passive formation, 222–223
  - plurals, 141
  - pronouns, 141
  - revitalisation, 395
  - stress, 248
  - vowel lengthening, 28
  - vowels, 193
  - word order, 282
- Māori English, 316
- Māori Party, 506

Māori Pidgin English, 396
marketing
  linguists in, 46–47
Martha’s Vineyard, 331
Martin, George R. R., 52
Matsumoto, Yoshiko, 469
Maxim of Manner, 460–461
Maxim of Quality, 459, 463
Maxim of Quantity, 457–458
Maxim of Relation, 459
maxims
  contravening, 461–463
  of politeness, 466
Mayer, Mercer, 441
Mc prefixes, 98
McBride, Cammie, 570
McBurney, Samuel, 309
McCulloch, Gretchen, 476
McGregor, Bill, 291
MD transcription system, 191, 192
Meakins, Felicity, 25, 379, 401, 552
mean length of utterance (MLU), 560
meaning
  communicating, 449–451
  components, 155–157
  and context, 157–160
  described, 146–149
  extracting, 461–465
  inclusion, 155
  properties, 149–150
  shifts, 163–169
  why words shift, 167–169
medical settings, 48
medieval English, 397–398
Melbourne vowels, 213
men, 148–149
mental health
  and linguistics, 48
merits, 72
metacognitive skills, 537
metalanguage, 5
metalinguistic awareness, 537
metaphors, 161–162, 240, 462
metathesis, 225–226, 343–344
METEO (translation system), 591
metonymy, 162–163
metrolingualism, 380
Middle English, 338, 346, 397
migration, 392
mime, 11
mind-body effect, 48

minimal pairs, 229–230, 248
mishearing, 534
misunderstandings
  in courtrooms, 39–40
  and dialects, 305
  in medical settings, 48
  in questionnaires, 67
Mitchell, Alexander, 192, 482
mixed languages, 400
modal verbs, 304
models-in-(inter)action, 62
modernisation, 503
modesty maxim, 466
modification, 137–138
modifiers, 270, 280
  dangling, 277
mondegreens, 169, 203–205, 533
monolingual mindset, 378–379
monolingualism, 377, 537, 577
Montgomery, Lucy Maud, 452
Morey, Stephen, 32, 384
morphemes, 113–115
  allomorphs, 126–128
  bound, 119–120
  derivational, 121
  free, 118
  grammatical, 128–132
  lexical, 128–132
  productive, 139–140
  pronunciation, 243–245
  types of, 115–120
morphology, 113
  and the brain, 131–132
  changes, 347–351
  inflectional, 120
  processes, 137–139
morphs, 126, 243–245
mortality of words, 108–109
motherese, 551
motorisation, 168
mouth,
  parts of, 182
movement, 268
multi-dimensional analysis, 597
multilingualism, 401
Murray, Lee, 489

n sound, 345
name-calling, 91
name-dropping, 92
names, 89–93

  - changing, 92–93
  - commonisation, 101
  - curious sounds, 246
  - presidential-sounding, 91–92
  - sound symbolism, 205
- and taboos, 90 *See also* brand names
- Narrative Scoring Scheme (NSS), 439–440
- narratives
  - cultural aspects, 438–441
  - structures, 437–441
- narrowing, 164–165
- nasal sounds, 182, 223, 226, 228, 337–338
- nations
  - and languages, 505–507
- nativist approaches, 549
- natural classes, 241–243
- natural language processing, 52
- Natural Semantic Metalanguage (NSM), 156–157
- Na’vi language, 52
- Nayler, Benjamin Suggitt, 356
- near-minimal pairs, 230
- negation, 563
- neighbourhood effects, 534
- neologasm, 82, 326
- networks
  - dense, 360
  - social, 313, 360
- neuroimaging techniques, 525, 527–528
- neurolinguistic research, 71
- neurons, 522
- neurophysiology, 520
- new information, 413–415
- New Zealand English
  - borrowings from Māori, 29
  - collocations, 595
  - diphthongs, 194
  - flapping, 357–358
  - high rising tone or terminal (HRT), 314–315
  - regional variation, 309, 310
  - and standardisation, 504
  - syllables, 198
  - vowels, 190–191, 211
- Nibun, Y., 576
- nocebo effect, 48
- Nolan, Francis, 52
- non-basic clauses, 259
- non-referring expressions, 147
- non-rhotic dialects, 184, 224–225
- non-standard sentences, 258
- non-standard variation, 41–42, 497–498
- Noonan, Peggy, 163

North Semitic, 251
noun-noun compounds, 96
noun phrases, 133, 270, 273, 277, 284
nouns, 132–134
    conversions, 101
    to verbs, 101, 134, 138
    with verb forms, 134
NS (Nihon Shuwa, Japanese Sign Language), 9
nucleus, 198
number, 123
number marking, 132, 133 *See also* plurals

O'Neil, Wayne, 252
objects, 260, 286
obligatoriness, 286
observation, 62
observational approaches, 549
observational data, 61–63
observer's paradox, 61, 64, 73, 471
obsolescence, 108
Ochs, Elinor, 60
Okrand, Marc, 52
Old Bailey Corpus, 59
Old English, 84, 248, 338, 341, 346, 351, 397
    cognates, 365, 366
    nasal sounds, 337–338
    word order, 351
Old Norse, 365
Oliver, John, 91–92
One New Zealand Foundation, 506
one-word stage, 561, 563
online communication, 432
online identities, 328–330
onomancy, 91
onomatopoeia, 7
onset, 198
oral sounds, 182
order of words, 13–14
oronyms, 203, 254, 533
orthographies, 249–252, 500–501
over-extension, 557
overt prestige, 485
Oxford English Dictionary, 16

Pacific and Regional Archive for Digital Sources in Endangered Cultures (PARADISEC), 53
palatal sounds
    consonants, 180
palatalisation, 227–228, 340
palato-alveolar consonants, 180
Palmerston English, 315

Pama-Nyungan language. *See* Gurindji
Pan American, 39–40, 209
Papua New Guinea, 500 *See also* Mali; Tok Pisin
paradigms, 14, 122–126
paralinguistic features, 202
paraphrasia, 542
Parseltongue language, 52
parses, 536
  rule-based, 585–588
part of speech (POS) tagging, 585
participant observation, 62
participants
  portrayals of, 74–75
  relationships with, 75
parts of speech, 132–136
Pasifika English, 316
passives, 419–420
passivisation, 273
pauses, 530
Pawley, Andy, 311
Pennsylvania German, 23, 353, 482, 500
  cross-linguistic influences, 386, 389, 394
  diglossia, 381
  morphological processes, 144
and silence, 471
perfect aspect, 125
performance, 18
performatives, 456
perlocutionary acts, 456, 457
Perrott, Deb, 48
person, 124
personal pronouns, 14
perspicuous, 460
Petersen, David J., 52
phonemes, 217, 219–222
  and child language acquisition, 556
  determining, 229–240
  and metaphors, 240
  natural classes, 241–243
phonemic paraphrasia, 542
phones, 217–222
phonesthemes, 171–172
phonetic changes, 344–346
phonetic similarity, 239–240
phonetic symbolism, 205–206
phoneticians, 39
phonetics, 178, 217, 220
  acoustic, 206–213
  articulatory, 178–179
  consonants, 179–187
  and spelling, 195–197

  - vowels, 187–195
- phonological adaptations, 554
- phonological aspects, 84–85
- phonological changes, 344–346
- phonological interference, 220
- phonological markers, 318
- phonological processes, 222–229
  - languages other than English, 253
- phonological rules, 589
  - writing, 245
- phonologically conditioned, 127, 244
- phonology, 217, 220
  - acquisition of, 554–556, 575
  - and careers, 220
  - changes in, 337–346, 359
- phonotactic constraints, 246–247
- phonotactics, 339
- phrasal category, 284
- phrase structure rules, 277–281, 585
- phrases, 266–269
- physiological aspects
  - simplicity, 359–360
  - speech production, 178–179
- pidgin languages, 396
- Pinker, Stephen, 83, 99
- pitch, 179, 200, 202
  - and babies, 551
- pitch traces, 201
- Pitjantjatjara, 89
- placebo effect, 48
- Placek, David, 47
- places of articulation
  - consonants, 179–181
- Plain English movement, 39, 497
- Planigale, Mark, 459
- Plant and Animal (Ethnobiology) Projects, 45–46
- Player, Vivienne, 314
- plosives, 182–183
- plural markings, 130–131
- pluralism, 506
- plurals, 126–127, 133
- Polari, 324
- politeness, 465–468
  - and cultural differences, 469–473
- political correctness, 492–493
- *politician*, 166
- polysemous words, 535
- polysemy, 153
- polysynthetic languages, 86, 348
- Ponceau, Peter Stephen Du, 86
- portrayals of participants, 74–75

positive (clauses), 258–259
Possible Word Constraint, 533
postalveolar sounds
    approximants, 184
    consonants, 180, 183
power, 74–75, 468
pragmatics, 449
predicates, 259, 260
predicative complements, 287
Prefect, Ford, 19
prefixes, 98, 99, 116–117
prejudices,
    linguistic, 37, 480–483
prepositional phrases, 278
prepositions, 274
prescriptive approaches, 15–17, 496
presentational markers, 318
prestige, 356, 485
    and standardisation, 306
presuppositions, 463, 464
prevarication, 6
priming, 526
principle of end-focus, 415
principle of end-weight, 415
principle of front focus, 417–418
printing, 4, 500
Prinzivalli, Paul, 39–40
productive morphemes, 139–140
productivity, 4
professional services
    linguists in, 49–51
progressive tense, 125
pronouns
    and directness, 466
    personal, 14, 124, 285
    and political correctness, 492–493
    reflexive, 425–427
prosodic features, 247–248
prosodic phonology, 245–248
prosody, 197–202
    and child language acquisition, 554
Proto Germanic, 364
    nasal sounds, 337–338
    palatalisation, 340
Proto-Indo-European (PIE), 366–369
psycholinguistic research, 71, 525
psycholinguists, 57
Pullum, Geoff, 170
puns, 152 *See also* jokes
purism, 486–495

qualitative analysis, 69, 70–71
quality.
  of sounds, 247
  of voice, 202
    See also Maxim of Quality
quantitative analysis, 68, 69, 71
quantity. *See* Maxim of Quantity
Queen Elizabeth II, 210
questionnaires, 66–67
questions, 563

r linking, 224–225
r sound, 180, 184, 224–225, 343, 344
ranking impositions, 468
reading, 526, 535
  child development, 567–569, 570, 577
rebus principle, 250
recast, 552
recipes, 438
recording
  conversations, 73
recursion, 280–281
Red Riding Hood tale, 413, 414–415, 417
Reddy, Michael, 451
reduction, 362
reduplication, 103–104
reference, 148 *See also* denotation
referring expressions, 147
reflexive pronouns, 425–427
reflexiveness, 5
regional dialects, 308–309
regional variation, 308–312
register systems, 201
registers, 317–320
relational antonyms, 151
relations
  grammatical, 273–277
  lexical, 150–155
    See also Maxim of Relation
relationships
  and research partnerships, 75
relative clauses, 263
relative power, 468
relativiser, 264
relevance theory, 459
repetition, 362–363
representation, 74–75
representational gestures., 473
research
  advocacy, 75
  approaches, 549–550

  - and benefits, 72
  - empowering, 75
  - ethical, 72–74, 75
  - language variation, 331–332
- and risks, 72
- research partnerships, 75
- respect, 72
- revival of languages, 392, 499–500
- rheme, 416–417
- rhotic dialects, 184, 224–225
- rhyme, 198, 245, 554
- rhyming slang, 199
- Richards, Ann, 162
- Rickford, John R., 71
- right-dislocation, 423
- Rillington Place murders, 322
- risks, 72
- Roden, Colin, 283
- Rodriguez Louro, Celeste, 59, 315
- roots, 115
  - bound, 118–119
  - free, 118
- round-trip translation, 592
- rules, 13, 18
- runic alphabet, 251
- Russian
  - aspect, 30
  - borrowings into English, 30
  - business, 159
  - cases, 274
  - colours, 519
  - described, 29–30
  - inflection, 122, 123, 130
  - palatalisation, 227
  - and politeness, 470
  - take, 115
- Russian speakers,
  - learning English, 220

- *s* (plural), 126, 132, 133
  - pronunciations, 132
- *s* (possession), 132
- safety in the workplace, 50
- Salzman, Michael, 392
- sample size, 65
- Sanskrit, 366
- Sapir, Edward, 516
- Sapir-Whorf hypothesis, 516
- sarcasm, 462
- Saussure, Ferdinand de, 355
- Saville-Troike, M., 439

Scandinavian
    borrowings into English, 106
Schembri, A., 98, 311, 313, 425
Schleicher, August, 368
Schmidt, Johannes, 369
Schreyer, Christine, 52
schwa, 127, 189
Schwarzenegger, Arnold, 453
scientific discourse, 430
Scots English, 304
Sebba, Mark, 501
second-language acquisition (SLA), 554, 571–572
    developmental patterns, 572
    impact of first language, 572–574
    success factors, 574
secret languages, 245–246
segmentation, 553, 554
selection
    variation, 504
semantic analysis, 585–588
semantic aspects, 83–84
semantic changes, 163–169
semantic fields, 150, 155, 323
semantic halos, 166
semantic paraphrasia, 542
semantic primes, 156–157
semantic roles, 273–277, 295
semantically troubling sentences, 14, 258
semantics, 146
semiotics, 6
Semitic languages
    vowel modification, 138
senses, 147
    changes, 163–169
sentences
    categories, 258
    and comprehension, 536
    order of words, 13–14
    types of, 14, 261–266, 289–292, 450
Severin, Allie, 66–67
sexual promiscuity in words, 148–149
Sharples, Pita, 506
Shaw, George Bernard, 219
shifts, 165–166
shortenings, 95–96
Siegel, Jeff, 41–42, 498
Sigley, R., 61, 77
sign language
    and gestures, 474–475
sign languages, 10–11
    acquisition of, 557

- described, 21
- numbers, 21
- regional variation, 311–312
  - *See also* Auslan

signified, 7, 8

signifiers, 7, 8

signs
- arbitrariness of, 8–10
- described, 6–8
- iconic, 7
- indexical, 8
- linguistic, 9
- symbolic, 8

silence, 446, 470–472

Silva, Tricia Cleland, 429

simple (clauses), 258–259

simple sentences, 261

simplicity, 359–360

Simpson, Homer, 118

Simpson, Jane, 400

situational variation, 317–320

skin conductance responses, 328

slang
- slanguage, 324–328
- and swearing, 324

slant brackets, 196, 217

slashes, 196, 217
- use of long, 235

Slavic languages, 29 *See also* Russian

slips of the ear or brain, 169–171

slips of the tongue, 88

Smith, Philip, 403

sn words, 172

sneeze, 172

social constructionist, 549

social context, 19–20, 355–358

social dialects, 312

social distance, 467

social factors, 360–362

social inclusion, 495–498

social markers, 313

social networking sites, 328

social networks, 313, 360

social variables, 313

social variation, 312–317

socially isolated children without language, 554

sociolects, 312

sociolinguistics, 16

sociolinguists, 57
- in education, 41
- in legal system, 39

sociophonetics, 210
sound changes, 211–213
sound correspondence, 364
sound patterns, 93–107
sound recordings, 61
sound symbolism, 146, 205–206
sounds
  addition of, 223–224
  in connected speech, 222–229
  consonants, 179–181, 182
  devoicing, 342
  length of, 247
  modification, 226–229
  perception, 211
  and spelling, 195–197
  to symbols, 196
  visualising, 206–209
    See also phonetics
Southern Standard British English (SSBE), 192
Southland English, 310
spaces
  and compounds, 97
Spanish speakers,
  learning English, 221
speaker profiling, 39
SPEAKING (mnemonic), 442
specialist language, 317–320
specific language impairments (SLI), 541, 551
spectrograms, 206, 208, 583
speech
  identifying segments, 584
  sounds in connected, 222–229
  telegraphic, 562–564
and writing, 13, 195–197, 431–433 *See also* female speech
speech acts, 456–457
speech communities, 302, 313, 330, 331–332, 360–362
speech disorders, 131–132
speech errors, 530
speech patterns, 356–357
speech production, 532
  physiological aspects, 178–179
speech recognition, 583–584
speech recognition technology, 52
speech synthesis, 590
speech therapy, 47–49, 220
spelling, 219, 252
  Americanisation of, 488
  and sounds, 195–197
spelling reforms, 219, 502
Sperber, Dan, 459
spoken language, 202–203

  - differs from writing, 11, 12, 195–197, 431–433
  - difficulties with understanding, 529–530
  - extracting meaning, 461–465
  - impoliteness, 465–468
  - most widely spoken, 408
  - paralinguistic features, 202
  - representation, 249–252
    See also conversations

Spooner, William Archibald, 531
Spoonerisms, 531–532
square brackets, 196, 217
stages in language acquisition, 554–556, 558, 564, 572
  - one-word, 561
  - three-word, 562
  - two-word, 561–562

Standard English, 136, 304, 335, 496
standardisation, 304–305, 503–504
  - and prestige, 306
  - stages, 503

Stebbins, Tonya, 28, 63, 133, 439, 458, 471, 572
stems, 116, 117
  - allomorphs, 127

Stephens, Richard, 68
stereotypes, 480–483, 494–495
Stirling, Lesley, 542
Stone, Harold, 587
stops, 182–183, 223, 230, 346
stress, 197–198, 248
  - verb and noun pairs, 138

stress-timed, 198
strings, 13
Stroop effect, 525
structuralism, 355
study designs, 65
study of language, 57
style guides, 490–491
styles, 320–328
subjects, 259, 260, 284–285
subordinate clauses, 270, 284
subordinates, 129
subordination, 263
subreption, 168
substitution test, 267, 271
sub-stratum, 388
suffixes, 98, 99, 117
Sullivan, Karen, 162
Sumarian script, 250
Summer Institute of Linguistics (SIL), 25
superliterate varieties, 265
suppletion, 139
suprasegmental information, 197, 245 *See also* prosody

survival of language, 492
swearing, 67–68, 324–328
    and taboos, 68
Swedish, 364
Swift, Jonathan, 356
syllabaries, 11
syllabic consonants, 199
syllabic systems, 249
syllable-timed, 198
syllables, 197–200
    and child language acquisition, 554
    parts of, 198
    structure, 245–246
symbols
    International Phonetic Alphabet (IPA), 186
    sounds to, 196
    transcription, 433–435
    used in book, xxi, 234
sympathy maxim, 466
synaesthesia, 162
synapses, 523
synchronic perspective, 20, 355, 356
synecdoche, 163
synonymy, 150–151
syntactic analysis, 585–588
syntactic markers, 319
syntax, 257
    and authorship, 272–273
    changes, 351–355
synthesise speech sounds, 588–591

t sound, 235, 237–238, 357
taboo words, 526, 527, 538–539
taboos, 65, 168
    and names, 90
    and swearing, 68, 327
tact maxim, 466
Tagalog, 118
Takeuchi, Masae, 63
Tannen, Deborah, 470
tapping, 218
Tasmanian Vernacular English, 311
Tayul, Julius, 28
teachers, 40–42
technical writers, 43
technology
    linguists in, 52–53
telegraphic speech, 562–564
tense, 124, 125
tense vowels, 190
terminology

  changes, 327
texts (mobile phone), 12, 70–71
theme, 416–417
theory of mind, 541–542
there-constructions, 420
thesauri, 148, 151
thought
  and language, 517–519
three-word stage, 562
Thurlow, Crispin, 70
Tibeto-Burman languages. *See* Turung
tip of the tongue (TOT) phenomenon, 532
Tiwi, 348
toddlers, 551, 554, 556, 562, 564
Tok Pisin, 398, 498
  as baby talk, 398
  and Mali, 382–383
Tolkien, J .R .R., 51–52
tone, 201–202
tone languages, 32, 201–202, 248
tongue twisters, 228
top–down models, 528
topic, 416–417
Tourette's Syndrome (TS), 540
Tourism Australia, 60
trademarks, 39, 221–222
traditional transmission, 6
transcription, 218, 436–437
  broad, 218
  HCE system of, 191, 192
  MD system, 191, 192
  narrow, 218
  phonemic, 196
  phonetic, 196, 218
transcription systems
  speech, 433–435
transfers, 385–386
transition relevance place, 436
transitivity, 282–289
translanguaging, 383, 384
translation, 403–405
  machine, 406, 591–592
transmission of language, 6
tree diagrams, 260, 262, 263, 266, 267, 269
Trench, Richard Chenevix, 17
Turing, Alan, 582
Turkic language, 501
Turkish
  morphological processes, 144
  vowels, 254
turn-taking, 436–437

Turung
  described, 31–32
  inflection, 130
  information structure, 444
  isolating, 347
  plural markings, 130–131
  stops, 230
  tone, 248
  transitivity, 294
  vowels, 193
Twain, Mark, 124, 275–276
two-word stage, 561–562, 563
typological classification, 281

umlauts, 228, 342
under-extension, 557
understatement, 168
UNESCO, 45
ungrammatical sentences, 14, 258
Universal Declaration of Human Rights, 496
universal grammar, 549
Unserdeutsch, 400
unwritten languages, 83, 229
Uptalk, 200 *see also* high rising tone or terminal (HRT)
urbanisation, 393
user-related differences
  regional variation, 308–312
  situational variation, 317–320
  social variation, 312–317
U-shaped performance, 572
utterances. *See* spoken language
uvular sounds
  approximants, 184
  consonants, 180

v sound, 181
values, 484
  and variation, 484
Valyrian languages, 52
variation
  age-related, 312, 313–316
  English, 300–303
  ethnic-related, 316–317
  free, 237–239
  and languages, 306
  non-standard, 41–42, 497–498
  regional, 308–312
  researching, 331–332
  selection, 503
  situational, 317–320
  social, 312–317

  - across space, 302
  - in speech patterns, 356–357
  - across time, 20–21, 302
  - and values, 484
    - *See also* standardisation
- vector space models., 597
- velar sounds, 226, 345, 346
  - consonants, 180
  - stops, 182
- velum, 182
- verb phrases, 277, 280
- verbal hygiene, 486–495
- verbs, 134–135
  - conversions, 101
  - inflectional forms, 120
  - to nouns, 101, 133, 138
  - with noun forms, 134
- and subjects, 284, 285
- vernacularisation, 506
- Vernon, Edward, 165
- verse (verb), 100, 372
- Victoria University of Wellington, 50
- video recording, 61
- Vienna-Oxford International Corpus of English, 593, 594
- Vietnamese, 201
- visualising sound, 206–209
- vocabulary
  - acquisition of, 556–558, 575
  - differences, 304
- vocal fry, 202
- vocal tract, 178
- vocalisations, 554
- voice
  - comparisons, 39
  - paralinguistic features, 202
- voiceless consonants, 181, 227
- voicing, 181
- voicing assimilation, 227
- vowel harmony, 228, 340–343
- vowel
  - lengthening, 28
- vowels, 187–195
  - addition of sounds, 224
  - Australian and New Zealand English, 190–192
  - and child language acquisition, 554
  - deletion of sounds, 222
  - diphthongs, 193–195
  - in other languages, 192–193
  - length of, 247
  - lengthening, 8
  - and lips, 188, 189

  modification, 137–138
  quality, 189

and spelling, 195
Waletzky, J., 437
Wardhaugh, Ronald, 37
wave model, 369
weeds, 371
Wernicke’s aphasia, 131, 543
Wernicke’s area, 521
White, E. B., 491
Whorf, Benjamin Lee, 516
Wierzbicka, Anna, 156, 455
Wigglesworth, G., 576
Williams, Cara Penry, 480
Willoughby, Louisa, 42
Wilson, Aidan, 213
Wilson, Deidre, 459
women, 148–149 *See also* female speech
word boundaries, 83
word classes, 132–136, 585
word embeddings, 597
word identification, 533–535
word order, 274, 289–290
  types of, 281–282, 351–353
word recognition, 584
words
  clipped, 95–96
  creation, 82, 93–107, 326
  defining, 82–86, 113
  and emotions, 68, 328, 441, 538–539
  force of, 456–457
  grammatical, 87
  grammatical properties, 85
  lexical, 86–87, 128
  mind-body effect, 48
  mortality, 108–109
  order of, 13–14
  phonological aspects, 84–85
  semantic aspects, 83–84
workplace
  language in, 50
  safety in, 50
Wren, Melissa, 488
writing
  child development, 569
  and communication, 43
  development, 250–252
  phonological rules, 245
  and speech, 11, 12, 13, 195–197, 431–433
  ways to write words, 11

writing systems. *See* orthographies
written language, 43, 326–327 *See also* online identities
Wug test, 559
Wyld, Henry Cecil, 527

yeah-no, 62, 430–431
Yiddish, 339
    word order, 290
Yoda's syntax, 268
Yoruba, 202
Young, Amanda, 468
youse, 484–485
Yupik, 86
    morphemes, 116

Zachrisson, R. E., 219
Zijlstra, Maria, 41–42
Zwicky, Anne, 106
Zwicky, Arnold, 106